SEX CHROMOSOME EVOLUTION IN FELIDAE AND ITS POTENTIAL EFFECTS ON MEIOTIC DRIVE AND HYBRID STERILITY

A Dissertation

by

WESLEY ALLEN BRASHEAR

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	William J. Murphy
Committee Members,	Terje Raudsepp
	Paul B. Samollow
	Christopher M. Seabury
Interdisciplinary Faculty Chair,	David W. Threadgill

May 2019

Major Subject: Genetics

ABSTRACT

The main motivation for this dissertation research was to gain a better understanding of mammalian sex chromosome evolution and its contribution to reproductive isolation between species. We first conducted a comparative study of the feline Y chromosome, beginning by sequencing approximately 10% (~2.4 MB) of the ampliconic region of the domestic cat Y chromosome. This allowed us to identify the major ampliconic gene families and repetitive units, and also provided a reference sequence resource for comparative analyses across the family. Using a combination of fluorescence *in-situ* hybridization, qPCR, and whole genome sequencing data from 20 different cat species, we found that Y ampliconic gene families are highly conserved across the family. These gene families remain conserved despite drastic changes in copy number and large structural rearrangements between species. The degree of conservation of the ampliconic region of the felid Y chromosome demonstrated the longevity with which these regions can be maintained, and provided additional evidence refuting predictions of the eventual demise of the mammalian Y chromosome. We then conducted a comparative analysis of four high-quality mammalian X chromosome assemblies, including human, pig, mouse, and the domestic cat, in order to assess patterns of amplicon evolution across Mammalia. Our results demonstrated that ampliconic regions were enriched for lineage-specific gene gains in all four species, but that a majority of these regions appear to have evolved in the common ancestor of eutherian mammals. We also determined that the mouse X chromosome is an outlier in the number of gene gains and losses, as well as the amount of X-linked ampliconic

sequence. Evolutionary breakpoints in linkage between the derived mouse X and the ancestral eutherian X chromosome gene order appear to be enriched for ampliconic regions. We conclude that these regions act as hot spots for lineage-specific gene gains, as well as large-scale structural rearrangements. Our findings have refined and altered conventional ideas of sex chromosome evolution across Mammalia, and provide a novel understanding of how ampliconic regions evolve. We demonstrated that X and Y chromosome ampliconic regions have been maintained over long periods of evolutionary time, and may have conserved functional roles.

DEDICATION

To my son, James.

ACKNOWLEDGEMENTS

I would first like to thank my mentor Dr. William Murphy for the unending support and patience he provided throughout my training as a graduate student. He is a truly wonderful advisor, teacher, scientist, and friend. I hope to one day be able to emulate his approach to mentoring graduate students as he has cultivated an amazing environment for nascent scientists to grow and expand their knowledge of the field and improve their critical thinking. I would also like to thank my committee members: Paul Samollow, Chris Seabury, and Terje Raudsepp. They have all provided unique insights into my research, and our discussions have greatly improved my understanding of biology and evolution, as well as the final draft of this dissertation.

I would also like to thank the current and former members of the Murphy Lab: Kevin Bredemeyer, Dr. Nicole Foley, Dr. Gang Li, Dr. Victor Mason, and Dr. Alex Meyers. They have all put up with my loud outbursts and obnoxious singing, and have each been wonderful friends and colleagues. This dissertation would not have been possible without their daily support and encouragement.

Finally, I would like to thank my family for their continued support, patience, and love throughout my tenure as a graduate student. My parents, Jimmy and Cindy Brashear, have always encouraged me to strive and to pursue whatever interests or paths I feel would make me the happiest and most content. They are both loving role models and I am truly grateful for all that they have done and continued to do for me and my family. I would also like to thank my beautiful wife, Regina Brashear. I cannot put into words what your love and support mean to me so I will just say this: you are my Loml.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This dissertation was completed under the supervision and guidance of the following committee members: William J. Murphy (Advisor), Department of Veterinary Integrative Biosciences, Faculty of Interdisciplinary Program in Genetics; Terje Raudsepp, Department of Veterinary Integrative Biosciences, Faculty of Interdisciplinary Program in Genetics; Paul B. Samollow, Department of Veterinary Integrative Biosciences, Faculty of Interdisciplinary Program in Genetics; and Christopher M. Seabury, Department of Veterinary Pathobiology, Faculty of Interdisciplinary Program in Genetics. All fluorescence *in situ* hybridization work in the second chapter was conducted by Terje Raudsepp. All other work conducted for this dissertation was independently completed by Wesley Brashear, with significant input, guidance, and consideration from William J. Murphy.

Funding Sources

This work was funded by grants from the Morris Animal Foundation, the Texas A&M University College of Veterinary Medicine, and the Winn Feline Foundation.

NOMENCLATURE

DMI	Dobzhansky-Muller Incompatibility
FISH	Fluorescence *in situ* hybridization
MRCA	Most recent common ancestor
MSCI	Meiotic sex-chromosome inactivation
MSUC	Meiotic silencing of unsynapsed chromatin
MYA	Million years ago
MY	Million years
PAR	Pseudoautosomal region
XAR	X-added region
XCR	X-conserved region
YAR	Y-added region
YCR	Y-conserved region

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

Page

CHAPTER I

INTRODUCTION

1.1 Motivation

> *"There is grandeur in this view of life, with its several powers, having been originally breathed into a few forms or into one; and that, whilst this planet has gone cycling on according to the fixed law of gravity, from so simple a beginning endless forms most beautiful and most wonderful have been, and are being, evolved."*
>
> - Charles Darwin, 1859

> *"I hope you're not gonna' bust our chops on this one, Bill."*
>
> - Steve Zissou, 2004

Quite simply, the impetus for this research was to understand the genetic basis for reproductive isolation between species as a way to gain insight into the mechanisms and processes involved in species formation - the origin of "endless forms most beautiful and most wonderful." As is often the case with questions regarding the natural world, the answers to this particular question first require answers to seemingly tangential or unrelated queries. When we consider the phenomenon of reproductive isolation,

particularly in its manifestation as hybrid sterility or inviability, certain patterns become apparent. The first is an observation made by J.B.S. Haldane in 1922:

> *"When in the F1 offspring of two different animal races one sex is absent, rare, or sterile, that sex is the heterozygous sex."*

Haldane observed this pattern across a number of animal species, including butterflies, birds, and mammals, and the same pattern has been documented when examined across a wider range of taxa (Coyne 1992; Wu and Davis 1993; Laurie 1997; Coyne and Orr 2004).

These general observations lead to an immediate and simple hypothesis regarding the genetic nature of reproductive isolation: sex chromosome evolution is a (the) major contributor to hybrid sterility/inviability. Indeed, countless studies since Haldane have provided both direct and indirect evidence for the disproportionate contribution of sex chromosomes to the evolution reproductive isolation (Coyne and Orr 1989; Coyne and Orr 2004; Presgraves 2008; Schilthuizen et al. 2011; Coyne 2018; Irwin 2018; Presgraves 2018). In fact, studies into reproductive isolation have been so prolific that shared genetic mechanisms or characteristics contributing to hybrid sterility/inviability have been uncovered and a number of well-tested hypotheses regarding their nature have been put forth (discussed below).

However, as a large number of these studies have been conducted using a very small number of organisms (e.g. *Mus*, *Drosophila*), it is useful to expand the

phylogenetic scope of these investigations to help inform any general conclusions made regarding the evolution of reproductive isolation. Fortunately, advancements in genomic technologies have led to a recent burst of studies investigating this phenomenon in non-model organisms (Noor and Feder 2006; Wolf et al. 2010; Ekblom and Galindo 2011; Nosil and Feder 2012; Ellegren 2014; Seehausen et al. 2014). These studies typically use whole genome and transcriptome sequencing data from populations to assess patterns of genomic divergence between species (or for pairs of taxa across a continuum of divergence), investigate patterns of introgression (i.e. clinal analyses), identify signatures of selection, map loci that are putatively involved in the phenotype (e.g. hybrid sterility), and evaluate differences in gene expression levels between parental species and hybrid offspring (Seehausen et al. 2014).

The cat family Felidae presents an ideal resource to conduct these types of analyses and to investigate the evolution of reproductive isolation. Interspecifc hybridization occurs naturally or through human mediation across much of this family, and several hybrids are regularly propagated between several species as pet breeds, providing a large hybrid population from which to sample (Gray 1972; Gershony et al. 2014; Davis et al. 2015). Davis et al. (2015) used two such interspecies cat breeds to conduct genome-wide association analyses to identify candidate regions involved in hybrid male sterility, and transcriptome data from hybrid testes to assess differential expression between parental species and sterile hybrid males. This study identified several autosomal and X-linked loci/regions associated with this phenotype and also revealed a chromosome-wide pattern of overexpression of genes from the X

chromosome in hybrid male testes. Unfortunately, regions associated with hybrid sterility along the X chromosome were large and typically associated with highly complex, repetitive regions of the chromosome. Furthermore, the contribution of the Y chromosome was not investigated as markers from the male-specific portion of the genome were not included on the 63K single-nucleotide polymorphism array that was used for this study.

In order to fine map the X-linked loci involved in reproductive isolation within Felidae, and to investigate the possible contributions of the Y chromosome to hybrid sterility, a better understanding of the genetic architecture and gene content of these chromosomes is required. In order to address these limitations, the research described in this dissertation sought to gain a better understanding of sex chromosome evolution by sequencing and assembling the complex regions of the domestic cat X and Y chromosomes, and by using the resultant data to help elucidate patterns of sex chromosome evolution across mammals that might contribute to the large role sex chromosomes play in reproductive isolation.

1.2 Sex chromosome evolution

The genetic determination of sex is a process that has independently evolved in a diverse array of eukaryotic taxa, including mammals, birds, reptiles, insects, and some flowering plants (Bull 1983; Charlesworth 1996; Rice 1996). This process, regardless of which sex is heterogametic (e.g., XY males in mammals, ZW females in birds), typically results in similar evolutionary trajectories for the proto-sex chromosomes (Charlesworth

and Charlesworth 2000; Charlesworth et al. 2005; Ellegren 2011). In the established consensus model, the process begins when a sex-determining allele arises on one homologue of an autosomal pair of chromosomes. Secondly, either through pre-existing linkage and/or recruitment through transposition, loci that are beneficial to the determined sex and are deleterious to the opposite sex (i.e. sexually antagonistic) begin to accumulate around the sex-determining locus. This results in selective pressure to suppress recombination within this region to maintain linkage between the sex-determining locus and alleles at neighboring loci that favor the determined sex. (Charlesworth et al. 2005; Bachtrog 2006, 2008; Ellegren 2011) (Fig. 1.1).

The suppression of recombination can be accomplished in two ways: (1) genetic modifiers that reduce the frequency of crossovers, or (2) genomic rearrangement of the region, typically through inversions (Brooks 1988; Charlesworth et al. 2005). In the most common models of sex chromosome evolution, the latter of the two alternatives is invoked as it has been demonstrated in a variety of taxa, including mammals, birds, and plants (Lahn and Page 1999; Handley et al. 2004; Hobza et al. 2007). The cessation of recombination after chromosomal inversion within the heterogametic sex allows for the independent evolution of the chromosome harboring the sex-determining allele, but also reduces the efficacy of natural selection in the removal of deleterious mutations. This results in the degeneration of the proto-Y or proto-W chromosome, with the accumulation of repetitive DNA and the loss of function in ancestral homologous genes (Bachtrog 2013; Hughes and Page 2015; Graves 2016; Wright et al. 2016). This leads to the rapid loss of homomorphy between the neo-sex chromosomes, and eventually results

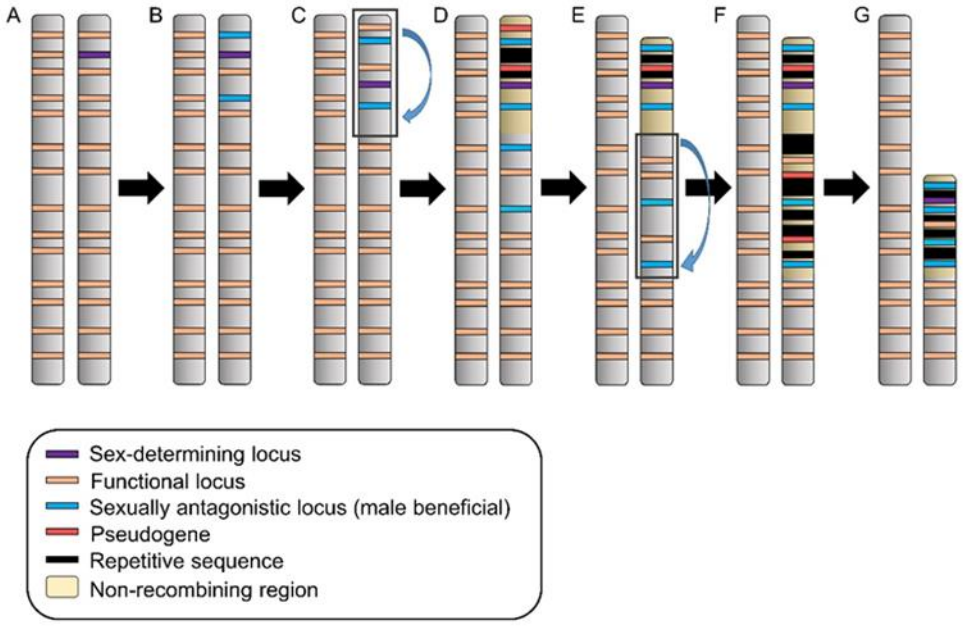


Figure 1.1. Model for the evolution of heteromorphic sex chromosomes. *(A)* A sex-determining locus is acquired on one copy of originally autosomal chromosomes. *(B)* Sexually antagonistic alleles (e.g. male-beneficial alleles deleterious to females) accumulate around the sex-determining locus. *(C)* The sexually antagonistic alleles cause natural selection to favor suppression of recombination within this region as to maintain linkage between them and the sex-determining locus. This suppression is achieved through chromosomal inversion. *(D)* Lack of recombination within the inverted region of the chromosome allows for the buildup of repetitive sequences and the pseudogenization of genes within this region through processes such as Muller's ratchet. Additional sexually antagonistic alleles become linked to the non-recombining region of the chromosome *(E)* Natural selection again favors the suppression of recombination between the additional sexually antagonistic alleles and the sex-determining region leading to the fixation of another chromosomal inversion. *(F)* Lack of recombination further degrades the non-recombining region leading to additional repetitive sequence, loss of sequence content, and more loss-of-function mutations. *(G)* Continued degradation due to lack of recombination results in heteromorphic sex chromosomes over time. (Adapted from Bachtrog 2013 and Wright et al. 2016).

in the characteristic heteromorphic karyotype of "ancient" sex chromosomes (Bachtrog 2013; Wright et al. 2016) (Fig. 1.1).

While this model of sex chromosome evolution has been developed from studies across a range of taxa, one of the most well studied systems with "ancient" sex chromosomes is the class Mammalia. The mammalian sex chromosomes, specifically those of the subclass Theria (marsupials and placental mammals), arose from an ancestral pair of autosomes approximately 165-180 million years ago, after the divergence from Monotremata (Portzebowski et al. 2008; Veyrunes et al. 2008). Following the split of the placental mammals and marsupials, the sex chromosomes in Eutheria acquired additional sequence (Spencer et al. 1991a, 1991b; Watson et al. 1992; Graves 1995; Toder et al. 1997; Lahn and Page 1999). This translocation is referred to as the X/Y added region (XAR/YAR), as to distinguish it from the regions of the sex chromosomes that share homology with Metatheria, referred to as the X/Y conserved region (XCR/YCR) (Graves 2006). The unique modes of inheritance, as well as disparate selective pressures, drove the mammalian sex chromosomes down drastically different evolutionary paths (Bachtrog 2013; Wright et al. 2016); and, as such, discussion of their history and general biology is best served through independent review.

1.3 The mammalian Y chromosome

The mammalian Y chromosome is comprised of a suite of distinct sequence classes (Skaletsky et al. 2003). The pseudoautosomal region, or PAR, is the only portion of the Y that still undergoes recombination with the X chromosome and allows the two to pair during meiosis. This region is variable in size and distribution, with some Y

chromosomes containing PARs at both the proximal and distal ends of the chromosome. For example, the domestic cat Y contains a single PAR on the proximal end of the chromosome that is approximately six megabases (Mb) in length (Li et al. 2013), while the human Y contains two pseudoautosomal regions comprising approximately 4.6% of the total length of chromosome, 2.7 Mb in length at the proximal end and 0.33 Mb at the distal end (Skaletsky et al. 2003; Flaquer et al. 2008; Otto et al. 2011).

The X-degenerate region of the Y contains genes retained from the ancestral proto-sex chromosomes, but no longer undergoes recombination with the X. This region is relatively small as most of the original genes and sequence have been lost, and the genes retained on the Y in placental mammals typically comprise of a subset of approximately 18 ancestral genes found in the most recent common ancestor of all Eutheria (Bellott et al. 2014). These genes are widely expressed, dosage-sensitive, and enriched for transcription and translation regulators, targeting regions throughout the mammalian genome (Bellott et al. 2014; Cortez et al. 2014).

The male specific region of the Y chromosome also contains the far less conserved ampliconic regions. These regions are highly repetitive, typically contain amplified gene families with testis-specific expression, and can differ dramatically between even closely related species in both gene content and copy number (Lahn and Page 1997; Murphy et al. 2006; Hughes et al. 2010; Paria et al. 2011; Cortez et al. 2014; Soh et al. 2014; Ghenu et al. 2016; Oetjens et al. 2016). Ampliconic regions of the genome are defined as those that contain large multi-kilobase duplicated segments (>10kb) that exhibit >99% nucleotide identity (Mueller et al. 2013). This high level of

pairwise identity between duplicated segments makes their sequencing and assembly using standard second generation sequencing untenable, and in draft assemblies these regions are typically collapsed or misassembled (Eichler et al. 2004; Alkan et al. 2011; Huddleston et al. 2014; Chaisson et al. 2015; Khost et al. 2017). Until recently, ampliconic regions on the sex chromosomes could only be assembled using a time and labor intensive Sanger-based clone sequencing approach (SHIMS – single haplotype iterative mapping) developed by the David Page Laboratory (Skaletsky et al. 2003; Hughes et al. 2010, 2012; Mueller et al. 2013; Soh et al. 2014). Given the cost and difficulty taken to accurately assemble these regions, the nucleotide sequence of the ampliconic portions of sex chromosomes are only sequenced in a small number of mammals: human (Skaletsky et al. 2003; Mueller et al. 2013), chimpanzee (Hughes et al. 2010), rhesus macaque (Hughes et al. 2012), and mouse (Mueller et al 2013; Soh et al. 2014). However, recent advances in sequencing technology (e.g. Single Molecule, Real-Time sequencing from Pacific Biosciences) and the development of new clone-based assembly protocols have reduced the time and cost required to generate draft sequences of these regions, and additional assemblies of mammalian Y chromosomes are being published (Tomaszkiewicz et al. 2016; Bellot et al. 2018; Janečka et al. 2018).

Given the non-recombining nature of the male-specific region of the Y chromosome, it is subject to a variety of degradative processes, the understanding of which is critical when contemplating Y-chromosome evolution. One such degenerative mechanism is known as Muller's ratchet, which involves the fixation of deleterious alleles in a finite non-recombining population through the stochastic loss of mutation-

free chromosomes. Within a recombining background, such as the X chromosome, deleterious mutations can be purged through recombination to establish haplotypes without harmful alleles. Due to the lack of recombination along the MSY, this mechanism is not available to remove deleterious mutations and once fixed are irreversible (Charlesworth 1978; Bachtrog 2013).

Y chromosomes can also accumulate mutations through a process called genetic hitchhiking (Rice 1987; Bachtrog 2013). In this instance, a beneficial mutation arises but is in linkage with one or more deleterious alleles. If the increase in fitness from the beneficial allele outweighs the costs of the deleterious allele(s), then the fixation of that beneficial mutation will also result in the fixation of any linked deleterious alleles (Bachtrog 2013). Again, this problem is unique to the non-recombining MSY as recombination between homologous chromosomes can disrupt linkage between the beneficial allele and any linked harmful mutations.

Although the Y chromosome is unique in its vulnerability to degradation due to the lack of homologous recombination, the distribution and orientation of ampliconic regions is such that intrachromosomal recombination and gene conversion allow these regions to repair deleterious mutations through self-correction (Rozen et al. 2003; Lange et al. 2009). Within the ampliconic regions, the large segmental duplications are often arrayed in a palindromic orientation. This orientation allows the chromosome to fold in such a way that if a harmful allele or loss-of-function mutation arises in one of the duplications, the mutation can be repaired through copying the non-mutated allele from

the opposite arm of the palindrome (Fig. 1.2). This process has allowed for the maintenance of ampliconic gene families in spite of a lack of recombination.

1.4 The mammalian X chromosome

Whereas the evolutionary history of the mammalian Y chromosome has largely been the product of transmission in the absence of recombination, the X chromosome still undergoes recombination (albeit at a reduced rate relative to autosomes) in females. The X chromosome is, however, subject to a number of biological characteristics and processes that differentiate it from autosomes. The X chromosome is hemizygous in males (present in only a single copy) and therefore faces selective pressures that are unique when present within the heterogametic sex. For example, if a recessive allele should arise on the X chromosome, the hemizygous nature of that allele immediately exposes it to natural selection in the heterogametic sex (Charlesworth 1978, 1996). In the case of male-beneficial genes, over time this process would result in the accumulation of these genes on the X chromosome (masculinization) (Rice 1984; Bachtrog 2006). However, the X chromosome spends a majority of its time within females allowing for selection to favor the accumulation of alleles beneficial to females (feminization) (Rice 1984; Khil et al. 2004; Bachtrog 2006). These models of X chromosome evolution are confounded, however, by other evolutionary pressures, such as the need to maintain gene-expression dosage equality between males and females (Charlesworth 1996; Graves 2016) and the process of meiotic sex-chromosome inactivation in males (Turner 2007) (discussed below). As a result, the evolutionary history of the X chromosome is

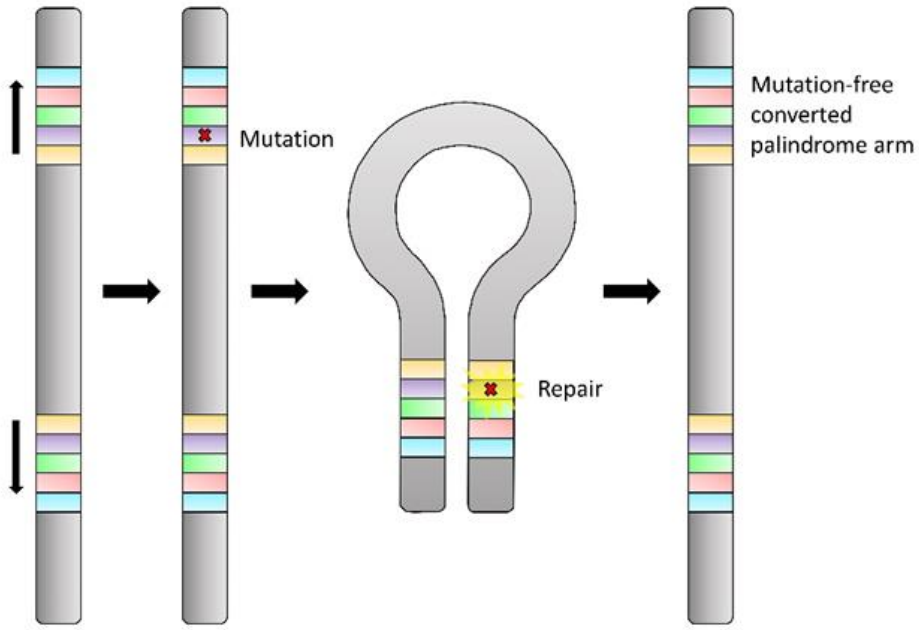


Figure 1.2. Mutation repair via intrachromosomal gene conversion. The colored blocks represent ampliconic duplications in a palindromic orientation.

quite convoluted and its study requires knowledge of processes that at first may seem tangential, but which undoubtedly affect the patterns that are ultimately observed.

As mentioned above, the evolution of heteromorphic sex chromosomes presents an obvious hurdle: the maintenance of gene expression levels between sex chromosomes and autosomes and between the homogametic and heterogametic sexes (Augui et al. 2011; Deng et al. 2014; Finestra and Gribnau 2017; Jégu et al. 2017). Within mammalian evolution, the impact of heteromorphic sex chromosomes on gene expression has been mitigated (or exploited) in two major ways. First, early within the

evolution of mammalian sex chromosomes there was extensive movement of genes on to and off of the X chromosome (Emerson et al. 2004; Vinckenbosch et al. 2006; Potrzebowski et al. 2008). Intronless retrocopies of X-linked genes moved to autosomes, presumably as a way to maintain dosage levels of housekeeping genes on the X chromosome during meiotic sex chromosome inactivation in males (discussed below) (Portzebowski et al. 2008). Conversely, the X chromosome has also recruited a disproportionate number genes when compared to the autosomes, with these genes frequently playing a role in the male germline (Wang et al. 2001; Lercher et al. 2003; Emerson et al. 2004).

Secondly, and perhaps the more studied consequence of the evolution of heteromorphic sex chromosomes, is the need to normalize expression between the homogametic and heterogametic sexes. This can be accomplished through increasing the expression of the X chromosome in the heterogametic sex or by decreasing expression of the same chromosome in the homogametic sex. In mammals, dosage compensation is achieved via X chromosome inactivation, a process that occurs in the somatic cells of females in which one copy of the X chromosome is silenced through epigenetic modifications (Ohno and Hauschka 1961; Lyon 1960; Augui et al. 2011). In 1967, Susumu Ohno predicted that the first step in the in the evolution of this process was the upregulation of X-linked genes to maintain levels of expression relative to the autosomes. This upregulation led to the overexpression of X-linked loci in females, which was resolved by the evolution of X chromosome inactivation (Ohno 1967, Graves 2016).

Ohno went on to predict that, after natural selection had resulted in the formation of the process of X chromosome inactivation, that any translocations from the X chromosome to autosomes would disrupt the dosage relationships between the X-linked and autosomal genes (Ohno 1967, Lyon 1992). This disruption would be deleterious and natural selection would select against these translocations, and thus the genes located on the X chromosome would not change between species throughout mammalian evolution (Ohno 1967, Lyon 1992). This hypothesis has become known as Ohno's Law, and has been well-supported by a number of studies (Prakash et al. 1996; Nesterova et al. 1998; Murphy et al. 1999; Watanabe et al. 1999; Band et al. 2000; Kuroiwa et al. 2001; Ross et al. 2003; Spriggs et al. 2003; Raudsepp et al. 2004; Delgado et al. 2009).

The selective pressures to retain the ancestral X-linked gene catalog are such that this conservation must be considered when contemplating the disproportionate role the X chromosome plays in reproductive isolation, as this conservation seems counter-intuitive when coupled with the high frequency with which hybrid sterility loci map to this chromosome. In fact, one might actually expect a dearth of these loci on the X given its considerable conservation across eutherians. However, in 2013, Mueller et al. conducted a study assessing Ohno's law between mouse and humans using assemblies and annotations that allowed for a finer scale investigation than previous comparative mapping approaches. They found that while a majority of X-linked genes within these two species followed Ohno's law, those that did not were enriched for function in male gametogenesis and had been independently acquired in each lineage. This suggests that these genes do not actually contradict Ohno's law, as the independent acquisition of

genes on the X chromosome can be independent from the retention of the ancestral gene repertoire. Of further interest, the genes not shared between these two species were most frequently found within ampliconic regions of the sex chromosome. Mueller et al. (2013) also points out that hybrid sterility loci in *Mus* hybrids often map near or within these amplicons.

Another source of constraint for X chromosome evolution would be the silencing of both sex chromosomes during male gametogenesis, a process shared across mammals called meiotic sex chromosome inactivation (MSCI). This process is an extension of meiotic silencing of unsynapsed chromatin (MSUC), and the failure of this process is known to be a factor in meiotic sterility (Turner et al. 2007). Interestingly, a failure of MSCI is a common phenotype in sterile mammalian hybrids studied to date, and may represent a common mechanism in the evolution of hybrid sterility (Good et al. 2010; Davis et al. 2015; Larson et al. 2016).

1.5 Sex chromosomes and speciation

The most common model of the evolution of hybrid sterility loci is known as that which came from the work of Bateson (1922), Dobzhansky (1936), and Muller (1942). This model suggests that reproductive barriers evolve due to a deleterious epistatic interaction between two or more loci, now called Dobzhansky-Muller incompatibilities (DMIs) (Wu and Ting, 2004). In this model, an interbreeding population splits into two separate lineages. In one lineage, a novel allele at one locus evolves and becomes fixed, while in the second lineage another novel allele at an interacting loci also undergoes

fixation within the population. While neither allele is deleterious in the lineage within which it evolved (otherwise natural selection would have purged the mutation from the population), when the two alleles interact within a hybrid genomic background, they lead to hybrid sterility or inviability (Fig. 1.3) (Wu and Ting, 2004).

While these deleterious epistatic interactions can occur at any location within the genome, a common observation is that these loci are enriched on the X chromosome. This pattern is so common in fact, that the "Large X-Effect", which refers to the disproportionately large role the X chromosome plays in hybrid sterility and inviability, is considered to be one of the "two rules of speciation", with Haldane's Rule being considered the other (Coyne and Orr, 1989; Coyne 1992; Turelli and Orr, 2000; Presgraves 2008; Presgraves 2018). The role of the X chromosome in hybrid sterility has been documented across a wide range of taxa, including *Drosophila* (Presgraves 2008), natural and laboratory populations of *Mus* (Dod et al. 2002; Larson et al. 2016), and felids (Davis et al. 2015).

Several hypotheses have been proposed to explain the phenomenon of the large X-effect, none of which are necessarily mutually exclusive. First, the dominance theory, suggests that the reduction in fitness observed in heterogametic hybrids (through means of fertility or viability) is a result of the deleterious allele being recessive in nature, but exposed as a result of its hemizygosity in the heterogametic offspring. The deleterious effect of this allele is shielded in the homogametic hybrids, as the dominant allele is present within their genomic background (Orr 1997; Turelli and Orr, 2000; Johnson 2010). A second hypothesis, dubbed the "faster-X effect", posits that loci contributing to

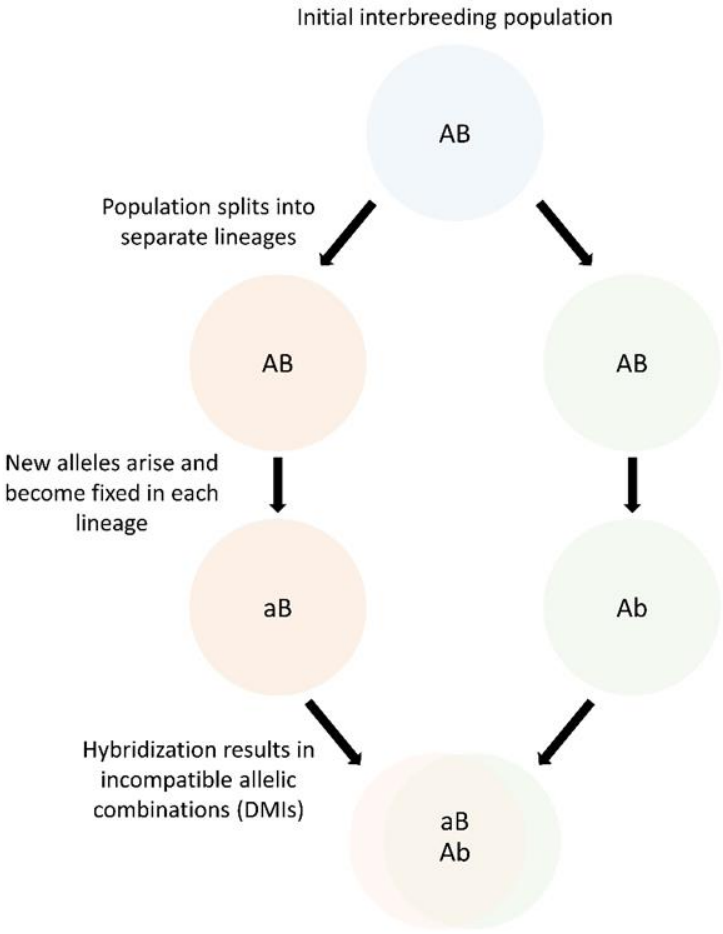


Figure 1.3. Model of the evolution of Dobzhansky-Muller incompatibilities. (Adapted from Payseur and Hoekstra, 2006)

reduced hybrid fitness arise and are enriched on the X because this chromosome is subject to a faster rate of adaptive evolution than are the autosomes, and, as such, more quickly accumulate possible incompatibilities (Charlesworth et al. 1987; Coyne and Orr 1989; Kirkpatrick and Hall, 2004; Vicoso and Charlesworth, 2006; Vicoso and Charlesworth, 2009; Mank et al. 2010; Connallon et al. 2012; Meisel and Connallon, 2013; Garrigan et al. 2014; Coyne 2018). Yet another hypothesis regarding the evolution

of hybrid incompatibilities is that male-specific genes evolve more quickly than female-specific or gender-neutral genes, and the rapidity with which these genes evolve is responsible for the effect of Haldane's rule in XY sex-determining systems (Schilthuizen et al. 2011). The faster rate of evolution is a result of male-specific or male-biased genes undergoing two types of sexual selection (both female-choice and competition between males) as female-biased genes are only subject to female choice (Wu and Davis, 1993; Wu et al. 1996; Schilthuizen et al. 2011).

Another, arguably more distinct, explanation for the observations encompassed by Haldane's Rule and the large X-effect, was independently proposed by Frank (1991) and Hurst and Pomiankowski (1991). This hypothesis posits that the evolution of meiotic drive systems, specifically those involving the favored transmission of one of the sex chromosomes, could contribute to the disproportionate effects of DMIs on the heterogametic sex (Patten 2017). This particular hypothesis seemed to originally garner little attention, but has since become more appreciated as some notable examples of this phenomenon seem to be correlated with hybrid sterility.

The contribution of sex chromosomes (particularly the X) to the evolution of reproductive isolation is demonstrably considerable. As such, a better understanding of the processes governing sex chromosome evolution is a critical and initial step in elucidating the molecular mechanisms and selective pressures that give rise to sterility phenotypes associated with reproductive isolation. With this dissertation, our goal was to expand our understanding of sex chromosome evolution by studying this phenomenon in the cat family Felidae. The benefit of this study system lies not only in the available

resources (e.g. high-quality genome assembly, BAC libraries, large sample population), but also in that extends the phylogenetic breadth in which sex chromosome evolution has been extensively studied. In Chapter II, we present data on the evolutionary history of the Y chromosome in Felidae, providing an initial draft assembly of the domestic cat Y and conducting comparative analyses across Felidae. We show that the ampliconic regions of the male-specific region of the Y can be stable over long evolutionary time spans, and, as such, that Y chromosome conservation is not restricted to the single-copy, X-degenerate region. In Chapter III, we conduct a broader comparative analysis of X chromosome evolution, comparing the sex chromosomes of the human, domestic cat, pig, and mouse. We show that the ampliconic regions found within mammalian X chromosomes, thought to be largely independent in their origin and evolution, are largely derived from common ancestral eutherian ampliconic sequence. In addition, we show that these regions likely serve as hot spots for breakpoints for structural rearrangements found within the mouse X chromosome.

CHAPTER II

EVOLUTIONARY CONSERVATION OF Y CHROMOSOME AMPLICONIC GENE FAMILIES DESPITE EXTENSIVE STRUCTURAL VARIATION*

2.1 Introduction

Advances in DNA sequencing technologies, alongside the increasing affordability of large-scale sequencing projects, have dramatically expanded opportunities for interspecific comparative genomics (Lindblad-Toh et al. 2011; Zhang et al. 2014; Lamichhaney et al. 2015; Hu et al. 2017; Meadows and Lindblad-Toh 2017). However, standard short-read sequencing approaches still fail to accurately assemble repetitive portions of eukaryotic genomes, collapsing large segmental duplications and incorrectly reconstructing complex genomic architecture (Eichler et al. 2004; Alkan et al. 2011; Huddleston et al. 2014; Chaisson et al. 2015; Khost et al. 2017). Mammalian Y chromosomes, specifically the ampliconic regions of the male-specific region of the Y chromosome (MSY), are enriched with large, high-identity segmental duplications that are prone to complex structural arrangements (Skaletsky et al. 2003). These characteristics have prevented the construction of accurate draft or finished quality assemblies of MSY amplicons using standard NGS techniques (Carvalho and Clark

---

*Reprinted with permission from Brashear, W.A., Raudsepp, T, &Murphy, W.J. 2018. Evolutionary conservation of Y chromosome ampliconic gene families despite extensive structural variation. *Genome Research* **12**: 1841-1851 by Cold Spring Harbor Laboratory Press under the Creative Commons License.

2013; Skinner et al. 2016; Tomaszkiewicz et al. 2016), leading to their absence in all but a small number of genome assemblies (Skaletsky et al. 2003; Hughes et al. 2010; Hughes et al. 2012; Bellott et al. 2014; Soh et al. 2014). As a result, most interspecific comparative analyses have been restricted to the ancestral single copy region of the MSY (Li et al. 2013, Bellott et al. 2014, but see Murphy et al. 2006, Cortez et al. 2014), which in most organisms represents a relatively small proportion of the Y chromosome's total size. While such comparisons have illuminated the evolutionary trajectory of mammalian Y-chromosome gene degeneration across broad time scales (tens of millions of years), a small subset of ancestral single copy Y-linked genes are highly conserved and widely expressed (Bellott et al. 2014). Previous studies have thus failed to illustrate the evolutionary dynamics of the more rapidly evolving, ampliconic Y chromosome sequence within the timeframe of recent adaptive radiations.

Outside of the conserved single copy region of the eutherian MSY, the ampliconic regions can differ drastically in both gene content and copy number between closely related species, which may reflect differences in reproductive life-history traits (Hughes et al. 2010; Cortez et al. 2014; Soh et al. 2014). Ampliconic regions typically harbor massively amplified gene families that exhibit testis-specific or testis-biased expression, suggesting a primary function in male gametogenesis (Lahn and Page 1997; Murphy et al. 2006; Paria et al. 2011; Soh et al. 2014; Ghenu et al. 2016; Oetjens et al. 2016). Accordingly, Y-linked mutations can have drastic effects on male reproductive fitness, producing phenotypes such as spermatogenic failure (Kuroda-Kawaguchi et al. 2001; Repping et al. 2002), abnormal sperm morphology (Touré et al. 2004; Touré et al.

2005; Cocquet et al. 2012; Campbell and Nachman 2014), and sex ratio distortion (Cocquet et al. 2012; Helleu et al. 2015). Thus, given the Y chromosome's proclivity for accruing structural rearrangements, and the manifestation of male-biased sterility in interspecific mammalian hybrids (Haldane's Rule), it seems likely that ampliconic gene divergence has played a critical role in the speciation process (Meiklejohn and Tao 2010; Good 2012; Soh et al. 2014). To investigate this possibility, however, a better understanding of ampliconic MSY variation and its evolutionary patterns is needed across a larger sample of closely related species with characteristics that are amenable to dissecting the genetic basis of reproductive isolation.

The cat family, Felidae, is an ideal system to study MSY evolution and its biological consequences across a broad range of divergence times (Murphy et al. 2006; Pearks Wilkerson et al. 2008; Li et al. 2013; Davis et al. 2015). This family currently contains 41 species divided amongst eight lineages that emerged during the Late Miocene, ~12 million years ago (mya) (Johnson et al. 2006; Li et al. 2016; Kitchener et al. 2017). Each of the eight species complexes diverged within the past 3-6 million years (MY), presenting a diverse and parallel sample of radiations with which to examine a variety of evolutionary processes, most notably interspecific hybridization (Li et al. 2016). Additionally, important research tools and resources, such as BAC libraries and high-quality reference genome assemblies, already exist for the domestic cat and can be exploited across the family due to the high degree of synteny and sequence identity (Davis et al. 2009, Montague et al. 2014, Li et al. 2016). Finally, there are numerous examples of anthropogenically derived felid hybrids, including several derived from

crosses to the domestic cat, which are actively propagated as popular cat breeds (Gray 1972; Gershony et al. 2014; Davis et al. 2015). Some hybrid crosses exhibit sex-biased litters when F1s are backcrossed to non-hybrid parental species (Table B2.1), and F1 and early generation backcross males exhibit hybrid male sterility (Davis et al. 2015). Thus, the felid radiation is a powerful model system for investigating the underlying genetic basis for phenotypes such as sperm morphology and hybrid male sterility, including potential contributions from Y-linked components. Dissecting the genetic mechanisms behind male infertility within this system may shed light onto the etiology of teratospermia observed in wild felids (Wildt et al. 1988; Roelke et al. 1993; Barone et al. 1994; Pukazhenthi et al. 2001; Pukazhenthi et al. 2006; Johnson et al. 2010; Terrell et al. 2016), and could potentially inform conservation efforts of endangered or threatened felids.

Murphy et al. (2006) used direct cDNA selection and fluorescence in situ hybridization (FISH) to identify four novel gene families and a single ancestral gene family present within the ampliconic region of the domestic cat MSY, all of which exhibited testis-specific expression. These gene families included two that appeared to be unique to the felid lineage: *TETY1* and *CCDC71LY* (formerly FLJ36031Y). Both genes derive from autosomal homologues from chromosomes A3 and A2, respectively. In addition, they discovered an ampliconic expansion of an ancestral Y-linked gene, *CUL4BY*, which has since been shown to be multicopy within the dog and the horse (Paria et al. 2011). Another ancestral Y-linked gene, *CLDN34Y* (formerly *TETY2*), appears to be multicopy exclusively within Carnivora (Li et al. 2013). A fifth gene,

TSPY1, is multicopy in nearly all mammalian Y chromosomes examined (Bellott et al. 2014; Cortez et al. 2014), and is highly amplified across the domestic cat Y. The small portion (<1Mb) of the domestic cat multicopy region that was previously sequenced (Li et al. 2013) contained additional gene orthologs found in other mammals, including *HSFY* (multicopy in primates, pigs, and cattle) and *RPS4Y* (multicopy in primates) (Cortez et al. 2014).

Although the transcriptional content of the ampliconic region of the domestic cat MSY has been previously explored, the actual DNA sequence, structure, and regulatory architecture has not been determined. Sequencing and assembling these regions is important to unveil heretofore undiscovered gene repertoires, enable structural variant analyses, and aid in comparative sequence analysis with closely related species. Furthermore, interspecific comparisons of gene content and copy number can reveal candidates for the causative genetic mechanisms behind such phenomena as hybrid male sterility and sex ratio distortion (Qvarnstrom and Bailey 2008; Good 2012; Mueller et al. 2013; Paudel et al. 2015). Here we used a multi-pronged approach that combines BAC clone-based finishing, qPCR, FISH, and whole genome sequence analyses, to reveal the overall pattern of ancestral and lineage-specific Y chromosome gene loss/gain events and structural rearrangements within the cat family. Our results show that the ampliconic regions became massively amplified prior to the divergence of modern Felidae ~15 Mya, and have retained an unexpectedly high degree of conservation of gene content, yet show extensive copy number variation and structural rearrangements, across the family. These findings have important implications for the future study of reproductive isolation

within Felidae, as well as our understanding of Y chromosome evolution across Mammalia.

2.2 Results

2.2.1 Ampliconic BAC clone sequencing and assembly

We sequenced 48 randomly selected BAC clones from a candidate pool of over 500 positive clones identified in a screen of half of the RPCI-86 domestic cat clone filters, using overgo probes designed from five domestic cat ampliconic MSY gene transcripts: *CCDC71LY*, *CLND34Y*, *CUL4BY*, *TETY1* and *TSPY1* (Murphy et al. 2006). FISH analyses from Murphy et al. (2006) revealed that the majority of ampliconic genes were dispersed along the whole length of the chrY long arm, suggesting it is composed of an array of similar repeat units. Clones were randomly selected, and included a suite of different autorad signal intensities. Six clones were sequenced using both the PacBio RSII and Illumina MiSeq platforms, while the remaining 42 clones were sequenced only on the Illumina MiSeq, as we primarily sought to test the adequacy of this more cost-effective short-read approach. Each clone that was sequenced with PacBio assembled into a circular contig. The average length of assembled clone inserts was 113,475 bps (S.D. 17,779), which falls well within the range of expected insert lengths. *De novo* assembly of the 48 Illumina libraries resulted in 20 clones that passed our established quality control (QC) metrics, including five of the six clones sequenced on both the MiSeq and PacBio platforms. The clones that did not pass QC failed at one or more stages of the assembly process. For example, in some clones, the initial assembly using

only the paired-end libraries was too fragmented to confidently scaffold using wgs mate pair libraries from the same cat (Fig. A2.1a). Secondly, visual inspection of the mate pair connections between contigs within a scaffold revealed inconsistencies in contig placement, insert size, and/or likely autosomal contamination within the assembled clone (a complete failure of mate pair reads to map over large segments of the clone) (Fig. A2.1b). Finally, we identified several clones that displayed disparities in read depth that suggested that portions of these sequences were misassembled or contained segmental duplications that were collapsed (Fig. A2.1c).

In total, 21 clones were accurately assembled with these approaches (19 clones via Illumina sequencing and 6 clones using PacBio sequencing with 5 clones being assembled by both methods), typically into single contigs or scaffolds (Supplemental Table B2.2), and represent ~2.4Mb (~10%) of the domestic cat ampliconic MSY. When we compared the MiSeq assemblies that passed QC to their PacBio counterparts, we found that there were no major disparities between the two sequencing approaches (99.9% average pairwise identity, Supplemental Figs. A2.2-A2.6), and a majority of these differences were small indels, frequently found in short tandem repeats. The sequence within the PacBio assemblies that correlated to the larger gaps present within the Illumina assemblies typically consisted of stretches of DNA that were either highly GC poor or GC rich, and the differences in length between assembly approaches were small (~1% average difference in length; Table B2.3).

2.2.2 Ampliconic Sequence Content

Analysis of approximately 10% of the estimated ampliconic sequence on the domestic cat MSY revealed that it is comprised of a mixture of sequence classes derived from several independent autosomal transposition events, as well as retention and expansion of ancestral X-Y orthologues (Fig. 2.1). This finding is consistent with previous FISH-based observations that several ampliconic transcripts "paint" the entire long-arm of the domestic cat Y chromosome (Murphy et al. 2006, Li et al. 2013). We identified three separate autosomal transpositions that appear to have occurred at different periods in the evolutionary history of the felid Y chromosome. We dated the timing of these transpositions using multi-species alignments and relaxed molecular clock approaches in conjunction with fossil constraints (see Methods). The oldest transposition is estimated to have occurred approximately 46 mya (24-66 mya 95% CI) (Figs. A2.7-A2.10), more than 30 my prior to the most recent common ancestor (MRCA) of all extant felids, and near the estimated time of the feliform radiation (Eizirik et al. 2010). This event copied ~83 kb from ChrA2 onto the ancestral feliform MSY, including the gene *CCDC71L*. The second autosomal transposition occurred 32 mya (18-53 mya 95% CI) (Fig. A2.11), also originating from ChrA2, 30 Mb upstream of *CCDC71L*. This 5.4 kb segment resides within an intron of the gene *SUMF1*. The last autosomal transposition occurred approximately 15 mya (11-23 mya 95% CI), which corresponds to the estimated age of the felid MRCA (Fig. A2.12). This ~100kb Y-linked segment originated from a ~219kb segment from the short arm of ChrA3 (Fig. 2.1). The transposed sequence includes the previously described ampliconic gene *TETY1*, as well

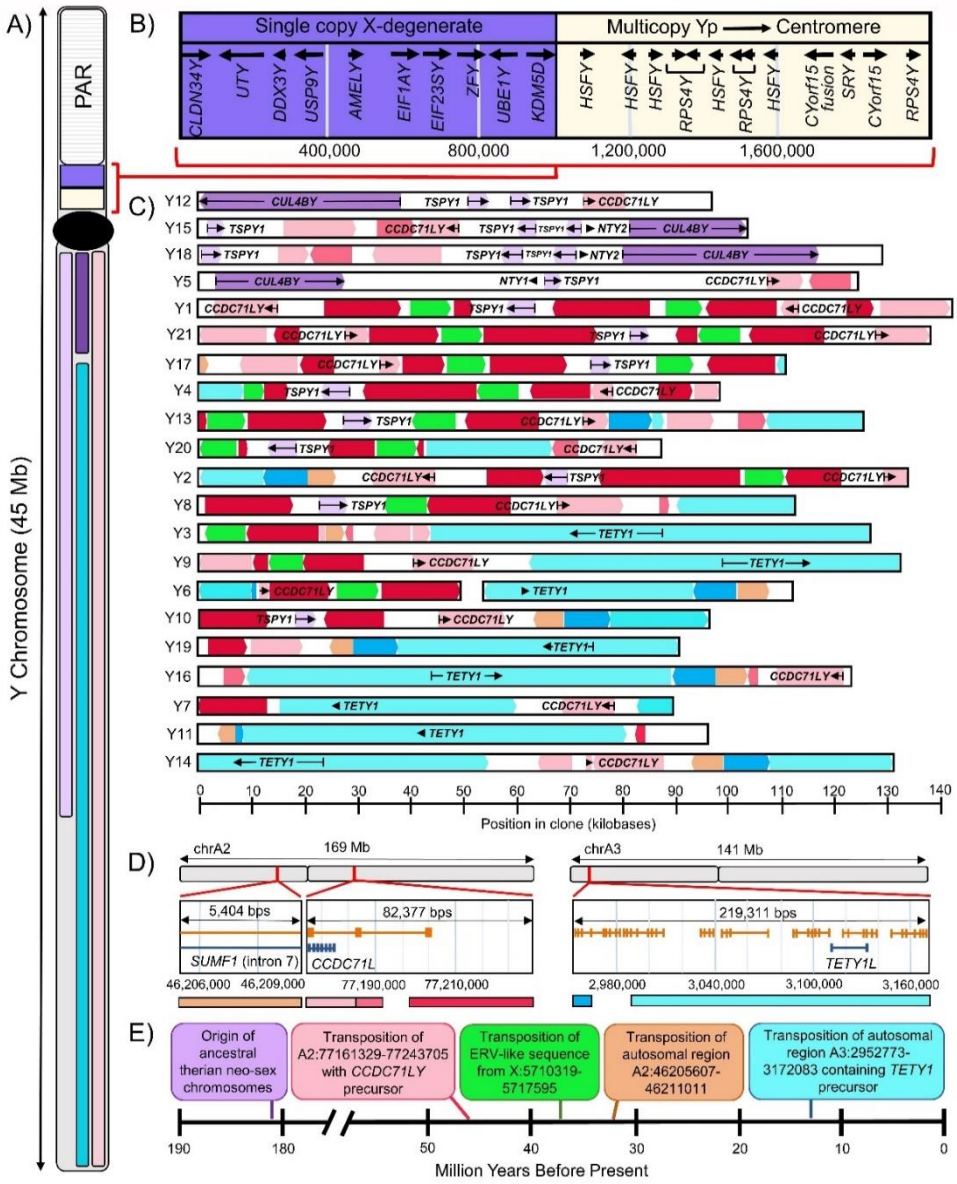


Figure 2.1. Domestic cat Y chromosome sequence content. (*A*) Ideogram of the domestic cat Y chromosome. The purple and beige regions of the short arm represent the single and multi-copy region of the Y chromosome sequenced by Li et al (2013). In the long arm, the colored blocks indicate the distribution of known ampliconic genes based on FISH analyses, with the corresponding genes being indicated in section 1c. It is important to note that the resolution of the FISH analysis did not allow us to determine the extent to which the ampliconic region extends across the centromere. (*B*) Annotation of the Y chromosome region as presented in Li et al. (2013). (*C*) Sequence content of 21 assembled BAC clones annotated with both transcripts (black arrows, text) and sequence origin (colors). Clones are placed in order by sequence content similarity and this order is not reflective of their chromosomal position. (*D*) Annotated view of autosomal progenitor sequences with the representative color for each transposed region displayed below. Annotated loci are depicted in blue and *in silico*-predicted loci are shown in orange. (*E*) Timeline of notable events in felid Y chromosome evolution.

as long stretches of flanking DNA. Although we estimated non-overlapping dates for the last two transpositions, the *SUMF1*-containing sequence originating from A2 is always found adjacent to the transposed material from chromosome A3 (Fig. 2.1).

Three ampliconic MSY genes arose from ancestral X-Y homologues: *CUL4BY*, *CLDN34Y*, and *TSPY1*. No single scaffold contained the entire coding sequence of *CUL4BY*. Three scaffolds, Y12, Y15, and Y18, had sequences that corresponded to the first five to nine exons of the published sequence of this gene, while scaffold Y5 contained sequence derived from the last four exons of the published sequence of *CUL4BY*. While these sequences could be pseudogenized copies, they were typically located at the ends of sequenced clones oriented in such a way that they may represent fragments of a complete gene sequence. It is worth mentioning that the *CUL4BY* transcripts constructed using the RNA-seq libraries revealed additional transcribed sequence within these paralogs. This could be due to true diversity in the *CUL4BY* variants or an artifact of the assembly process. We identified numerous intact copies of *TSPY1* (present in ~52% of sequenced clones), often with multiple genes per clone. We were unable to identify any additional copies of *CLDN34Y* within any of the assembled ampliconic scaffolds, which may be due to its low estimated copy number (Li et al. 2013).

We identified a modified endogenous retrovirus (ERV) derived from an X-linked sequence that has spread along the long arm of the Y chromosome (Fig. 2.1) after its initial retrotransposition approximately 20 mya (14-27 mya 95% CI) (Fig. A2.13). This ERV sequence is found on several autosomes, but is enriched on the X and Y

chromosomes ($\chi 2=283$, d.f. = 19, $p < 0.001$; Fig 2.2A-2.2C; Supplemental Table S4). Furthermore, the Y-linked copies of this sequence are highly diverged from their non-ChrY counterparts (Fig. 2.2A), yet retain fragments of the genes encoded by ERVs (e.g. gag, pol, pro, env). We identified multiple sequences within the ERV consensus sequence that could possibly act as promoters for nearby genes (Fig. 2B).

2.2.3 Ampliconic Y Transcripts

Using an RNA-seq library derived from domestic cat testis tissue, we identified several previously undescribed transcripts, as well as a large amount of previously unidentified diversity in felid Y-linked ampliconic gene families. Murphy et al. (2006) reported three versions of *TSPY1*: a244, b274, and c264. These three variants comprise two subgroups, with *TSPY1a* and *TSPY1c* sharing ~98% pairwise identity, while *TSPY1b* only shares ~64% identity with both *TSPY1a* and *TSPY1c*. We identified a large number of additional *TSPY1* variants, all of which fall into one of the two subgroups previously identified (Fig. 2.3). The *TSPY1b* subgroup clusters with *TSPY1* variants found in dog and human, while the subgroup containing *TSPY1a* and *TSPY1c* are more similar to the cow variant. This suggests that *TSPY1* was ancestrally multicopy but was lost independently in several lineages. The rodent lineage does not cluster closely with either of the two subgroups, although it should be noted that Tspy1 has been pseudogenized in some members of the mouse lineage (Schubert et al. 2000), and the reconstructed tree is unrooted.

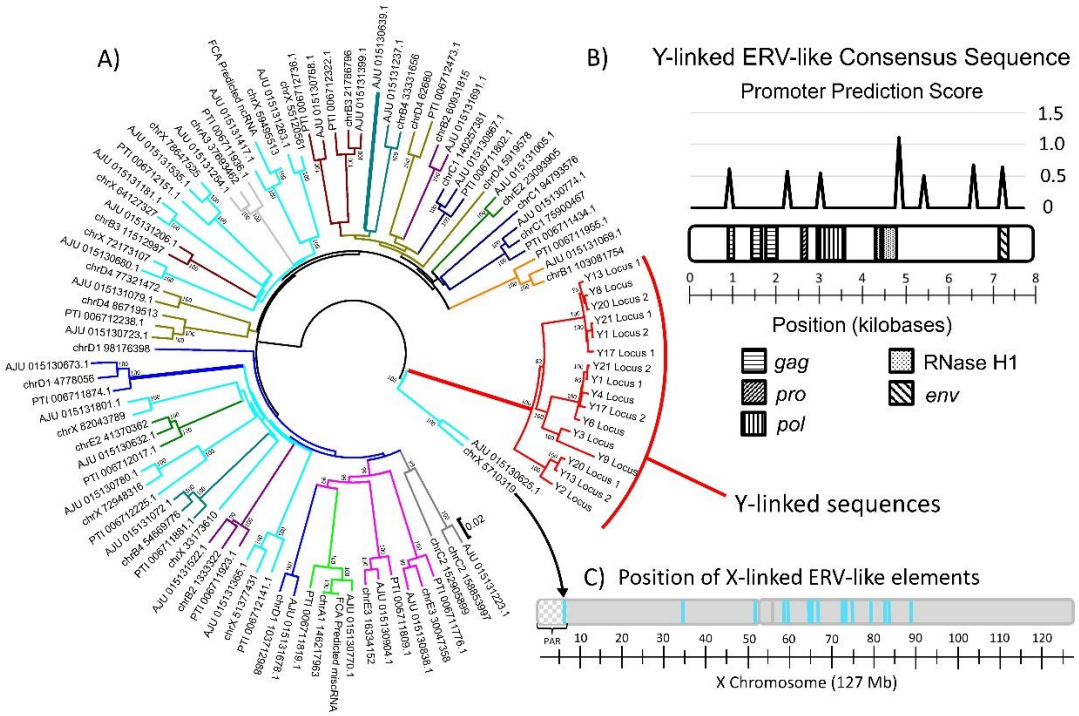


Figure 2.2. Paralogous endogenous retrovirus (ERV)-like elements are enriched on the X and Y Chromosomes. (*A*) Phylogenetic reconstruction of relationships between ERV-like elements identified within the domestic cat genome (labeled by chromosome and coordinates), cheetah (AJU), and Tiger (PTI). The branches are color-coded based on their predicted chromosome of origin. (*B*) Ideogram of consensus of Y-linked ERV-like sequences annotated with vestigial ERV genes (below) and the promoter prediction score (above) by position within the sequence. Scores below 0.5 can be ignored, while those ranging from 0.5-0.8 are considered marginal, between 0.8 and 1.0 as moderately likely, and any score above 1.0 is considered as highly likely. (*C*) Position of ERV-like sequences along the domestic cat X chromosome.

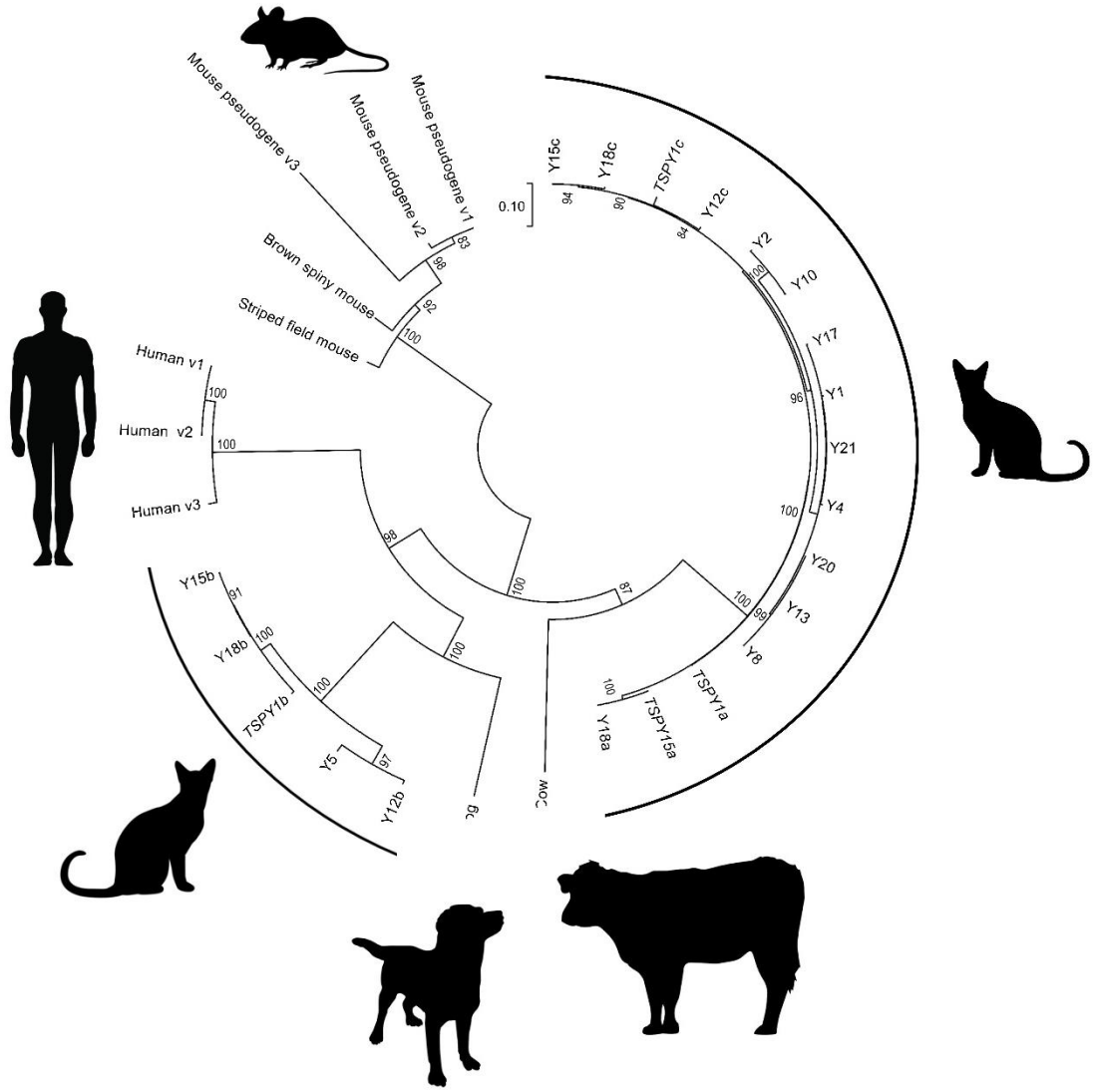


Figure 2.3. Phylogenetic reconstruction of the *TSPY1* variants published for the domestic cat, dog, human, cow, and several rodent species, as well as *TSPY1* transcript variants identified in this study (all clones starting with an uppercase "Y").

Skaletsky et al. (2003) reported an antisense open reading frame, *CYorf16*, on the human MSY that was transcribed from each *TSPY1* locus. An ortholog for this open reading frame was not found within any of the reconstructed cat *TSPY1* transcripts. We

did, however, identify five novel *TSPY1*-like open reading frames (ORFs) within a substantial number of the ampliconic scaffolds, ranging in length from 100-220 amino acids, and including both sense and antisense orientations (Fig. A2.14). Whether these candidate open reading frames undergo active transcription and translation remains to be determined.

The *CCDC71LY* ampliconic gene family contains a large number of previously undescribed transcript variants, all of which retain open reading frames in the sense direction ranging from 105 to 318 amino acids. Although the nucleotide sequence across this large family is moderately conserved (~93% pairwise identity), there is a large amount of sequence diversity among the different open reading frames in this gene family. We identified the presence of both sense and antisense ORFs (99 to 322 amino acids long; Fig. A2.15), wherein an array of frameshift and frameshift-recovering mutations has led to a diversity of putative peptide sequences comprised of different combinations of shared sequence motifs (Figs. A2.16-A2.18). While we have no data supporting the translation of these open reading frames, these results do display the presence of a large suite of potential proteins.

Finally, we identified two previously undocumented, testis-specific Y-linked transcripts: *NTY1* and *NTY2* (Novel Transcript on the Y). *NTY1* is a 611 bp sequence that was found on only one Y scaffold (Y5) and contains a single exon with an open reading frame of 183 amino acids. *NTY2* was found in two scaffolds, Y15 and Y18 (99.8% identical), is 597 bps in length and contains a single exon with an open reading frame of 163 amino acids. These loci appear to be related, but are highly divergent from one

another (~66% pairwise identity). Neither transcript shared significant similarity to any sequence in the NCBI nr database, nor did they contain any repetitive elements. Both *NTY1* and *NTY2* were transcribed at levels comparable to that of known ampliconic genes (Fig. A2.19). Interestingly, *NTY1* occurs in scaffold Y5, which contains the longest stretch of DNA (~108 kb) with no clear homology to any autosome or the X chromosome (Fig. 2.1).

2.2.4 Ampliconic structural evolution

We employed a variety of methods to examine MSY structural evolution across the entire cat family. We obtained qPCR and *in-silico*-based estimates of copy number variation for Y-linked ampliconic genes within and across species, and observed a large degree of gene family size variation for each gene examined, even within a single species (Tables 2.1-2.2, Tables B2.5-B2.6). We also determined the autosomal progenitor of one multicopy Y-linked gene, *CCDC71L*, has undergone expansion within the Pantherinae, including the leopard, lion, tiger, jaguar, clouded leopard, and snow leopard (Tables 2.1-2.2).

To visually validate the amplification and chromosomal distribution of ampliconic genes, we conducted fluorescence *in-situ* hybridization (FISH) with labeled cDNA probes from the domestic cat. Hybridization was performed on chromosomes of five species from different lineages within Felidae: the Asian golden cat (*Catopuma temminckii*; Asian golden cat lineage), tigrina (*Leopardus tigrinus*; Ocelot lineage),

bobcat (*Lynx rufus*; Lynx lineage), serval (*Profelis serval*; Caracal lineage), and the snow leopard (*Panthera uncia*; Panthera lineage).

FISH confirmed the high degree of conservation of gene content across felid Y chromosomes. However, despite the conservation of content, the Y chromosomes of different cat species differ drastically in structure (Fig. 2.4, Figs. A2.20-A2.24). For example, *TSPY1* appears to be expanded along a majority of Yq in both the Asian golden cat and bobcat, but it is restricted to the proximal region of tigrina Yq. Similarly, *CCDC71LY* is massively amplified along tigrina Yp, but along Yq of all other felid Y chromosomes that we analyzed (Li et al. 2013). While the FISH results support the conservation of known ampliconic genes across the felid lineage, it should be noted that some lineages of Felidae may have independently acquired novel Y-linked genes that would go undiscovered using this method. For instance, the cDNA hybridization signal observed for the tigrina was restricted to the distal and proximal portions of the short and long arm, respectively; and the putative gene content of the remainder of the long arm is unknown.

Table 2.1. Copy number for Y-linked ampliconic genes and the autosomal gene *CCDC71L* based on *in-silico* estimates made from mapping whole genome sequencing data. Estimates for *CCDC71L* are solely based on mapping results from females to ensure estimates were not skewed by highly divergent Y-linked sequences mapping to the autosomal progenitor. Samples for which estimates are denoted as "NA" are lacking WGS data for one sex from that species. Samples with an asterisk (*) are averages taken from multiple samples.

Species	***TSPY1ac***	***TSPY1b***	***CCDC71L***	***CCDC71LY***	***CUL4BY***	***TETY1***	***CLDN34Y***
Domestic cat	90*	44*	2*	247*	20*	127*	2*
Wild cat	95	48	NA	270	20	137	NA
Sand Cat	123	51	NA	231	16	162	15
Black footed cat	59	63	NA	46	19	25	10
Cheetah	89*	115*	NA	224*	7*	13*	8*
Lion	52	52	4	48	6	485	8
Leopard	NA	NA	5*	NA	NA	NA	NA
Tiger	25*	21	4	106*	7*	40*	4*
Snow leopard	NA	NA	5	NA	NA	NA	NA

Table 2.2. Copy number estimates for Y-linked ampliconic genes based on quantitative real-time polymerase chain reactions. Primers for *TETY1* and *CUL4BY* showed bias in efficiency with increasing phylogenetic distance from the domestic cat and are therefore not included. Samples with an asterisk (*) are averages taken from multiple samples.

	TSPY1ac	***TSPY1b***	***CCDC71L***	***CCDC71LY***	***CLDN34Y***
Domestic cat	96	88	1	124	3
Chinese desert cat	130	96	1	128	7
Black footed cat	65	71	1	86	10
Jungle cat	71	48	1	110	82
Asian leopard cat	136	72	1	37	38
Pallas cat	78	132	1	184	1
Iberian lynx	55	77	1	10	6
Lion	140*	164*	27*	118*	21*
Leopard	118	95	25	260	15
Jaguar	104*	73*	19*	38*	25*
Tiger	54	43	15	12	12
Snow leopard	53	49	25	61	46
Clouded leopard	50	114	18	99	2

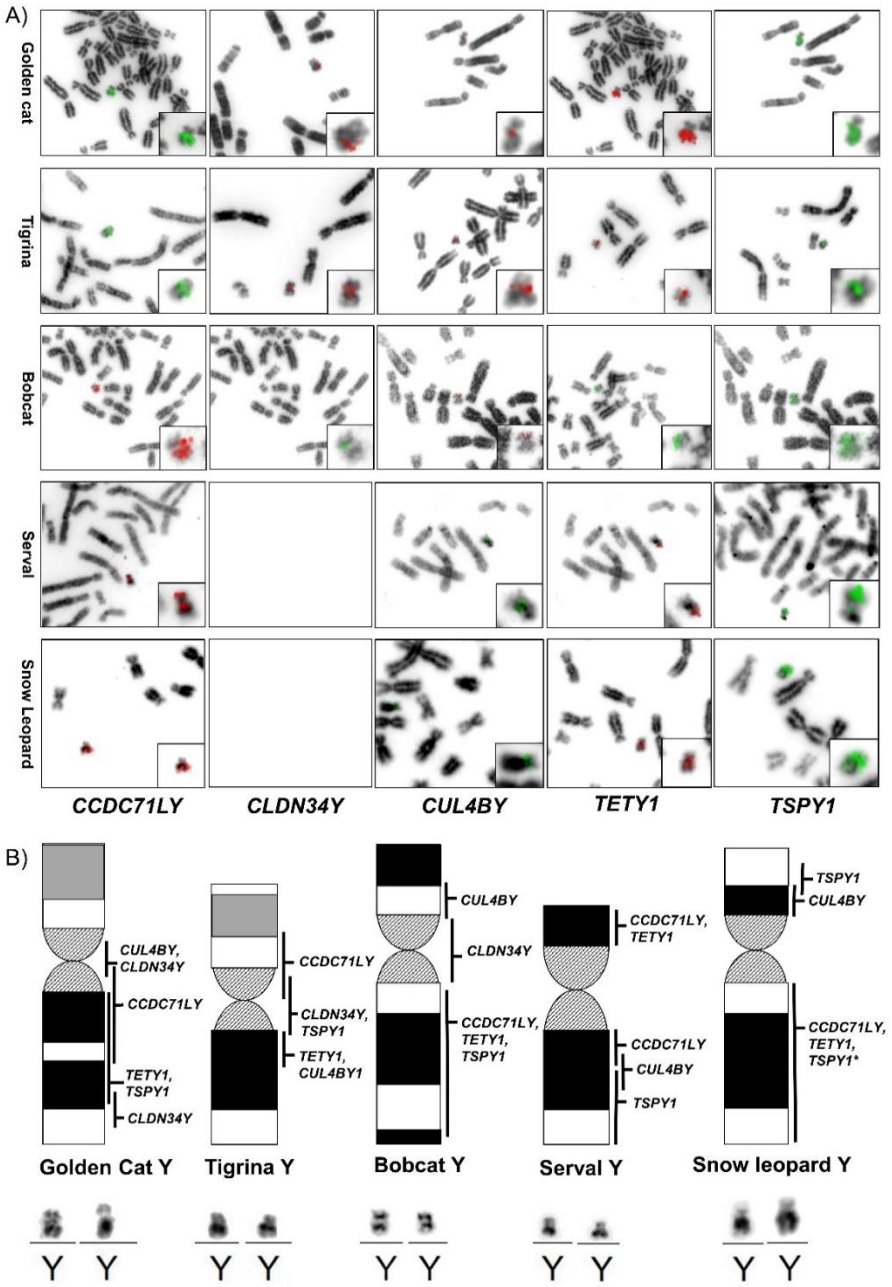


Figure 2.4. Distribution of ampliconic loci along the Y Chromosomes of five divergent felids. (A) Fluorescent in-situ hybridization with probes designed for known ampliconic genes in metaphase preparations of several felid species. (B) Ideograms for each species with the predicted distribution of each ampliconic Y-linked gene indicated. Note the long arm of the Tigrina Y contains a substantial amount of sequence in which no known ampliconic genes were identified. **TSPY1* appears to be present along the long arm of the snow leopard Y, but it should be noted that the signal intensity of this probe was not as strong as other ampliconic genes expanded throughout this region.

2.3 Discussion

2.3.1 Assembly strategy and results

We generated ~2.4 Mb (~10%) of randomly selected, ampliconic sequence from the domestic cat Y chromosome using two distinct sequencing platforms: Illumina MiSeq and PacBio RSII. While we were unable to confidently assemble every BAC clone sequence using our modified approach, we were able to accurately reconstruct many clones for a fraction of the cost that would have been required to conduct PacBio sequencing of every clone (i.e., high identity ampliconic clones would require PacBio barcoding). Regardless of the approach taken, our results affirm previous studies which concluded that accurate assembly of Y chromosome amplicons requires targeted sequencing approaches given their highly repetitive and structurally complex architecture (Hughes and Page 2014). These targeted strategies include BAC-based methods, or chromosome flow sorting and sequencing on several platforms, followed by extensive bioinformatic analyses and validation through bench work (Tomaszkiewicz et al. 2016). The increasing availability of Y-linked sequence information for a variety of species has beneficial implications for a variety of facets of biology. These resources can aid in elucidating historical and current population structures (Karmin et al. 2015; Hallast et al. 2016; Poznik et al. 2016; Jobling and Tyler-Smith 2017), allow for a more comprehensive understanding of male reproductive biology (Kuroda-Kawaguchi et al. 2001; Touré et al. 2004, 2005; Campbell et al. 2012), and provide a more complete substrate on which to investigate genome evolution (Murphy et al. 2006; Li et al. 2013). These factors are crucial to study the genomic consequences of sex ratio distorters,

recurrent invasion of selfish genetic elements on a non-recombining background, and reproductive isolation.

2.3.2 Remarkable ampliconic protein coding diversity

During spermatogenesis, specifically within meiotic spermatocytes and postmeiotic round spermatids, transcription across most of the genome is prolific, presumably due to permissive epigenetic regulation rather than playing any functional role (Soumillon et al. 2013); although it could be argued that spermatogenesis is a complex process requiring a large suite of disparate genes. Furthermore, lncRNA expression becomes increasingly testis-biased with decreasing evolutionary age (Necsulea et al. 2014), suggesting that novel gene evolution or gene family expansion is frequently associated with a gain of function within the context of the permissive environment of the testis transcriptome. Accordingly, we identified more than 50 testis-specific ampliconic gene family members, as well as two previously unidentified novel loci (*NTY1* and *NTY2*), and validated their specificity through analysis of RNA-seq data from a suite of domestic cat tissues (Fig. A2.19, Table B2.7). Further examination of this diverse array of variants within the context of their genomic architecture revealed a remarkable amount of putative protein coding ability. Specifically, within the largest two gene families, *TSPY1* and *CCDC71LY*, we observed 66 and 83 open reading frames, respectively, including both sense and antisense orientations (Figs, A2.14-A2.18). While these ORFs appeared to be relatively conserved across the *TSPY1* family, numerous small indels occurring across the different *CCDC71LY* variants has resulted in more

variable peptide combinations. The presence of such diversity suggests that the unique evolutionary processes that shape the Y (e.g. lack of recombination, higher mutation rate, smaller effective population size, drift, and male-limited transmission) provide (or in the least permit) a great deal of potentially functional variation.

2.3.3 Co-opting ERVs

Within the Y scaffolds we identified an expansion of a novel MSY sequence derived from an endogenous retrovirus (ERV). These ERV-like elements retained a long terminal repeat (LTR) as well as remnants of other retrovirus-encoded proteins (e.g., gag, pol). While this particular class of ERV is found elsewhere throughout the genome, it is notably enriched along the sex chromosomes (Fig. 2.2). Furthermore, this sequence does not appear to be randomly distributed throughout the assembled Y scaffolds, but was only found adjacent to a region that was transposed from ChrA2 ~46 mya, suggesting the segments were acquired together, and later amplified as a single unit. Specifically, the ERV co-localized with a region downstream of *CCDC71L* that contained neither annotated genes nor any Y-linked transcripts. The immediate ancestor of the Y-linked ERV-like elements is an ERV located within the pseudoautosomal region, suggesting a probable recombination-mediated mechanism by which the elements were acquired on the Y chromosome from its X chromosome predecessor.

Reduced recombination relative to the autosomes could account for the increased frequency of these potentially deleterious ERV-like insertions on the sex chromosomes, but the domestication of selfish genetic elements by the host to utilize their protein

coding potential or to regulate transcriptional activity has been documented by several studies (e.g., Werren 2011; McLaughlin and Malik 2017). Recently, Davis et al. (2017) demonstrated that LTRs from some mouse ERVs regulate the expression of many lncRNAs in both spermatocytes and spermatids. These findings suggest that the acquisition, retention, and amplification of this Y-linked feline ERV could serve to regulate *CCDC71LY*, and possibly other ampliconic Y-linked genes, during meiotic and post-meiotic stages of spermatogenesis.

2.3.4 Extensive copy number variation across Felidae

We observed substantial copy number variation for all ampliconic gene families across Felidae using both qPCR and *in-silico* based methods. This variation appears both within and between species, and is supported by differences in Y chromosome length observed in FISH analyses. The discrepancies between the qPCR and *in-silico* approaches highlights the difficulties encountered in studying complex ampliconic regions of mammalian genomes. The differences observed in this study are likely the result of multiple sources, both biological and methodological in nature. First, individual variation in Y chromosome length has been observed within species, and the individuals used for qPCR and the *in-silico* based approaches were different. This highlights the volatile course of Y chromosome evolution, and suggests strong selective pressure to retain these genes throughout the radiation of Felidae.

However, the methodological sources of the observed disparities should not go undiscussed. While the primers used for our qPCR experiment were based on multi-

species alignments of ampliconic loci, the variation observed within gene families within the domestic cat are highly suggestive of variation we were unable to capture in our primer design. For example, qPCR estimates for both *TSPY1ac* and *TSPY1b* were higher than those seen in the *in-silico* based approach, but the opposite pattern was observed in estimates of *CCDC71LY* copy number. This would be expected given the higher sequence similarity within *TSPY1ac* (98%) and TSPY1b (98.5%) than within *CCDC71LY* (~93%). Mapping short read data is likely more permissive of sequence divergence than qPCR, as discrepancies between primer and target sequence would lead to inefficiencies in PCR amplification. However, while the primers were designed to reduce the risk of cross-amplification, unknown variation within the autosomal progenitors across species creates the possibility co-amplification of autosomal and Y-linked copies.

Another issue we encountered when estimating copy number in-silico was mis-mapping of reads, specifically when mapping more distantly related species to the domestic cat reference genome. This problem can arise when divergence between species causes sequence diversity between orthologs and paralogs to overlap. We were able to mitigate these issues by utilizing both male and female samples to estimate copy number, allowing us to better ensure Y-linked reads were not mapping to paralogous autosomal genes and skewing results. The increasing quality and decreasing costs of long-read whole genome assembly will also help mitigate these issues in future studies that include high quality assemblies of each felid species.

2.3.5 Co-expansion of CCDC71L and CCDC71LY in Panthera

The autosomal progenitor of CCDC71LY, CCDC71L, has undergone amplification within Panthera (Tables 2.1- 2.2, Tables B2.5-B2.S6), perhaps in response to the amplification of *CCDC71LY* on ChrY. This may suggest the presence of an evolutionary arms race, in which the autosomal copy was amplified in order to suppress biased transmission of the Y. In support of this hypothesis, when female hybrid F1 ligers (offspring of a male lion x female tiger cross) are backcrossed to the lions, the litters display skewed sex ratios (Table B2.1). A similar process has been documented in *Mus*, in which mice deficient in copy number for the ampliconic Y-linked gene Sly produce female-biased litters and those with ampliconic X-linked *Slx*-deficient genotypes produce male-biased litters (Cocquet et al. 2012; Good 2012). In Panthera, however, the repressor of the Y-linked *CCDC71LY* would be autosomal and not sex linked, and, depending on the molecular mechanisms behind repression, could conceivably lead to male-biased litters produced by *CCDC71L* deficient males. A similar drive system has been found in Drosophila, but this involves the suppression of an X-linked sex-ratio distorter by an autosomal locus dubbed nmy (Tao et al. 2007). This suppressor evolved through a retrotransposition of the distorter itself and repression appears to be facilitated through sequence homology - a plausible mechanism in Panthera if *CCDC71LY* and *CCDC71L* do in fact constitute such a system.

Another intriguing possibility is that *CCDC71LY* and *CCDC71L* do not share an antagonistic relationship, but rather work cooperatively to regulate some particular function during reproduction. In humans, *CCDC71L* appears to be most highly

expressed within the fallopian tube (Fig. A2.25, DeLuca et al. 2012; Lonsdale et al. 2013; Carithers et al. 2015; Melé et al. 2015), while in the cat its expression appears to be highest in both the fallopian tube and testes (Fig. A2.26). These observations, taken together with the testis-specific expression of *CCDC71LY*, suggests the possibility of a functional relationship between the two gene families that is worthy of future investigation.

2.3.6 Y chromosome stability

Despite increasing evidence in support of the longevity and stability of ancestral single copy Y-linked genes throughout mammalian evolution, papers that espouse the Y chromosome's inevitable demise continue to be published (Graves 2016). Although the evolutionary forces that shape the mammalian Y are in some ways more drastic than other regions of the genome, our study, along with several others before, indicate that portions of the Y chromosome have not only remained stable, but most mammalian MSYs have accrued and amplified new genomic content (Murphy et al. 2006; Hughes et al. 2010; Hughes et al. 2012; Bellott et al. 2014; Soh et al. 2014). While the ancestral therian Y chromosomes did undergo an initial period of rapid degradation, the current model of gene loss on the Y predicts a non-linear loss of loci that has all but ceased in many species (Bachtrog 2008; Hughes et al. 2012; Bachtrog 2013). Hughes et al. (2012), for example, demonstrated that the rhesus macaque MSY has not undergone any gene loss in the last 25 million years, and gene loss along the human MSY has been restricted to the most recent evolutionary strata. Furthermore, Bellott et al. (2014) and Cortez et al.

(2014) demonstrated that ancestral Y-linked genes that were retained on the eutherian Y were not random, but were selectively retained due to dosage constraints and frequently maintain their ancestral protein functions.

Our findings suggest that the evolutionary stability of the Y is not limited to the more conserved single copy region, but can be extended to the ampliconic regions of the MSY. While human and chimpanzee MSYs differ drastically in both their gene repertoire and structure (Hughes et al. 2010), possibly due to differences in reproductive strategies of the two species, the gene content of felid Y chromosomes appears to be highly conserved across the family, although the relative copy number and chromosome structure appear to vary greatly between species. Reproductive strategies also vary greatly across the cat family (Macdonald and Loveridge 2010), therefore the apparent stability in ampliconic gene content is especially impressive given that the length of time that these gene families have been expanded and retained on the Y is far greater than the estimated age of the felid MRCA (Figs. A2.7-A2.13).

The conservation of ampliconic loci in the face of frequent structural rearrangements, processes like Muller's ratchet, and hitchhiking of deleterious alleles, suggests a high level of selective pressure to retain high copy numbers of these male-specific genes. This selective pressure enforces a type of chromosomal stability that differs from conventional ideas of genomic conservation (e.g., synteny), retaining certain functional Y-linked ampliconic genes while mitigating the erosive effects of replication without recombination. This is likely aided by the presence of palindromic sequences in

the ampliconic regions that allow for gene conversion (Rozen et al. 2003; Bachtrog 2013).

In summary, our analysis of Y chromosome evolution within the cat family illustrates the dynamic evolutionary forces that shape this chromosome. We provide evidence that genomic preservation and conservation is not limited to the shared (albeit depleted) X degenerate regions, but is extended to the ampliconic gene families and sequences that putatively play a large role in male reproductive success. Ampliconic genes are retained while the MSY is still serving as a dynamic substrate upon which novel genes and co-opted genomic elements undergo selection, most frequently in processes like male reproductive fitness, transmission distortion, and, perhaps, the establishment of nascent species boundaries.

2.4 Methods

2.4.1 Sequencing and assembly of the ampliconic Y

Clones containing Y chromosome fragments were identified by screening the male cat RPCI 86 BAC library using an equimolar pool of overgo probes derived from known Y-linked ampliconic genes: CCDC71LY, CUL4BY, TETY1, CLDN34Y, and TSPY1 (Murphy et al. 2006). Overgo probe design and filter hybridization methods were performed as previously described (Pearks Wilkerson et al. 2008). Selected clones were cultured and DNA was extracted using standard protocols. Extracted DNA was then sequenced on the PacBio RS II system using the P6-P4 chemistry and/or on the Illumina MiSeq platform using the 2 x 300 bp configuration. For Illumina sequencing,

extracted DNA was sheared to a target size of ~800 bps in order to increase the ability of the de novo assembler to span short (<1kb) repetitive segments. Mate pair libraries from whole genome DNA extractions from the same male cat were also utilized, with insert sizes of approximately 3kb and 8kb (SRX2376195 and SRX2376203, respectively). The mate pair libraries were processed using Nextclip v1.3.1 (Leggett et al. 2014) and mapped to the domestic cat 8.0 genome assembly (Montague et al. 2014) using default settings in BWA-MEM (Li and Durbin 2009). Mate pairs that did not align or that had a mapping quality < 30 were kept and represented those mate pairs most likely to be derived from the Y chromosome.

Sequenced Illumina libraries were trimmed using TrimGalore!, a wrapper tool around Cutadapt (Martin 2011), and were assessed for quality using FastQC. Trimmed libraries were mapped to the Escherichia coli K12 MG1655 genome using default settings in BWA-MEM (Li and Durbin 2009). Accurately aligned reads (mapping quality >= 30) were removed from the libraries. Filtered libraries were then split into subsets to target 100X coverage and each subset was assembled using A5-miseq (Coil et al. 2015) with default settings. The target coverage for each clone was adjusted to optimize assembly length and contiguity. Contigs were filtered for quality requiring an average per base phrap score of 75 or greater. Assembled sequences were then aligned back to the E. coli K12 MG1655 genome using BLAST (Altschul et al. 1990) to further remove any bacterial contamination. We then aligned the pTARBAC2 vector sequence to each assembly in order to identify and remove the vector from each sequenced clone.

After the assembly for each clone was quality filtered and had the vector sequence removed, we aligned the 3kb and 8kb mate pair libraries using Newbler version 2.9 requiring 40 bp overlap and 99% pairwise identity. This program was used in order to produce the output required for Consed (Gordon and Green 2013), which allowed us to easily visualize the alignments, read-depths, and mate pair connections for manual curation (Li et al. 2013). The order and orientation for scaffolding were produced by Consed and manually edited as needed. We then mapped the filtered paired end libraries back to the complete scaffold using the program Bowtie2 (Langmead et al. 2009) in Geneious (Kearse et al. 2012) with default standards for a final quality assessment. Mapping results were visually inspected for conspicuous anomalies in read depth coverage (e.g., the standard deviation of read depth was greater than the mean), and those without uniform coverage were not considered accurately assembled and were not used in downstream analyses.

Barcoded PacBio libraries were assembled using the PBcR pipeline with the Celera Assembler version 8.3rc2 (Koren et al. 2012; Koren et al. 2013; Berlin et al. 2015). Assemblies were circularized using minimus2 in AMOS (Treangen et al. 2011). The pTARBAC2 vector sequence was aligned to each circularized clone and clipped from the assembly. Clones for which we had both PacBio and Illumina data were aligned to assess the ability of our short read sequencing approach to accurately assemble these regions.

2.4.2 Sequence content of the ampliconic Y

We examined the sequence content of the assembled clones using a variety of tools in order to identify the origins, repeated units, and transcribed elements of the ampliconic Y. Repetitive elements were identified using the program RepeatMasker version open-4.0.5 (Smit et al. 2015). We used GMAP (Wu and Watanabe 2005) to identify the locations of previously described Y-linked ampliconic genes and to serve as a guide in identifying the origins of regions transposed from autosomal and/or X-linked progenitors. We iteratively aligned sections of assembled clones to the domestic cat genome assembly (Felis catus v8.0, which does not include the previous Y chromosome assembly) using BLAT (Kent 2002) to delineate the boundaries of transposed segments.

Once the shared constitutive segments from each clone were identified, we aligned them with the corresponding autosomal and/or X-linked sequences from the domestic cat, tiger (*Panthera tigris*), cheetah (*Acinonyx jubatus*), domestic dog (*Canis familiaris*), and panda (*Ailuropoda melanoleuca*) using the program MAFFT (Katoh and Standley 2013). Columns in the alignment matrix for which more than 50% of the sites were missing were removed and phylogenies and time trees were constructed using the programs RAxML (Stamatakis 2006) and MCMCTree (Yang 2007), respectively. Hard bound constraints for the base of Felidae were set at 10 and 24 million years for divergence time estimates (Johnson et al. 2006). Some Y-linked segments were not included in order to maximize alignment lengths for this analysis. Final alignment files are included in the supplemental materials.

In order to identify previously undescribed Y-linked transcripts, we created an RNA-seq library from testis tissue taken from a mature domestic cat. The sequenced library was iteratively mapped to the repeat-masked domestic cat genome, each time containing a single repeat-masked assembled clone as well as the domestic cat Y chromosome sequence published by Li et al. (2013), using the program STAR (Dobin et al. 2013). Transcripts were constructed and expression levels were estimated using Cufflinks (Trapnell et al. 2012). We downloaded RNA-seq libraries from the SRA database which were created from a variety of domestic cat tissues (Supplemental Table S7). These libraries were mapped to our augmented reference genome to characterize the tissue expression profiles of identified Y-linked transcripts.

After examination of the tissue expression profile *CCDC71L* in humans and noting the highest levels of expression in the fallopian tubes, we created two RNA-seq libraries from the same domestic cat fallopian tube sample. This allowed for technical replicates when we assessed the expression of this *CCDC71L* in the domestic cat for this tissue. Expression profiles for *CCDC71L* were quantified using the same protocol as above.

2.4.3 Estimating copy number of Y-linked ampliconic genes

We used two approaches to estimate copy number variation of Y-linked ampliconic genes for a large number of samples. We conducted qPCR on species for which we had genomic DNA using the relative comparative threshold cycle (CT) method on a Roche 480 LightCycler. We used primers based on multi-species

alignments for each ampliconic gene and for the single copy Y-linked gene *DDX3Y* to serve as a control (Table B2.8). Reactions were run in triplicate for each gene and the copy numbers were calculated using the 2-ΔΔct method (Schmittgen and Livak 2008).

To further validate the results obtained in the qPCR analyses, we estimated ampliconic gene copy numbers in-silico by analyzing changes in read depth coverage of WGS libraries mapped to the domestic cat reference genome (FelCat8.0) augmented with consensus sequences created from the assembled ampliconic regions of the domestic cat Y chromosome, including a single representative sequence for each Y-linked gene. These WGS libraries include several breeds of domestic cat, as well as a sampling of different species across the felid phylogeny, all downloaded from the SRA database (Table B2.6). The resulting alignments were then analyzed using CNVnator (Abyzov et al. 2011) to estimate the number of copies of each ampliconic gene across Felidae, using a bin size of 150 for all genes except *CLDN34Y* for which we used a bin size of 50 bps.

For the autosomal gene *CCDC71L* we used calls from CNVnator for the regions immediately flanking this gene. This was necessary as the domestic cat genome assembly contains a gap in the sequence within the annotation of this gene and, thus, reads were unable to accurately map to this stretch of the genome. We also made estimates solely based on short read libraries from female samples to avoid the possibility of overestimating copy number due to Y-derived paired end reads incorrectly mapping to this autosomal progenitor.

2.4.4 FISH

Fluorescence in-situ hybridizations were performed with biotin- and/or dioxigenin-labeled probes designed from cDNA sequences published by Murphy et al. (2006). These probes were hybridized to metaphase chromosomes from the Asian golden cat (Catopuma temminckii), tigrina (Leopardus tigrinus), bobcat (Lynx rufus), serval (Leptailurus serval), and the snow leopard (Panthera uncia) following established procedures (Raudsepp and Chowdhary 2008). A Zeiss Axioplan2 fluorescent microscope equipped with Cytovision/Genus V. 2.7 (Applied Imaging) was used to capture and analyze images from the metaphase spreads.

CHAPTER III

ANCESTRAL X CHROMOSOME AMPLICONS ACT AS HOT SPOTS FOR NOVEL GENES AND EVOLUTIONARY BREAKPOINTS

3.1 Introduction

The X chromosome is arguably the most evolutionarily unique and historically well-studied chromosome of the mammalian genome (Morgan and Bridges 1916; Lyon 1961; Ohno 1967; Penny et al. 1996; Ross et al. 2005). Perhaps one of the most interesting characteristics of the X chromosome is the remarkable degree of conservation in gene content across eutherian mammals, a property imposed by the need to maintain dosage relationships between X-linked genes and their autosomal counterparts (Ohno 1967; Lyon 1992). This pattern of conservation, unparalleled by any other chromosome within the mammalian genome, has been documented in numerous comparative studies (Nesterova et al. 1998; Murphy et al. 1999; Spriggs et al. 2003; Raudsepp et al. 2004; Delgado et al. 2009), although some exceptions have been noted (Palmer et al. 1995; Rugarli et al. 1995; Mueller et al. 2013).

In addition to the conservation of the ancestral X-linked gene repertoire, another remarkable feature of the mammalian X chromosome is the extent of linkage (i.e., gene order) conservation (*sensu* Nadeau 1989) displayed between phylogenetically distant lineages (Murphy et al. 1999; Quilter et al. 2002; Raudsepp et al. 2004; Delgado et al. 2009). Linkage conservation does not fall within the scope of Ohno's original prediction as intrachromosomal rearrangements presumably would not disrupt dosage levels to the

same degree as would interchromosomal translocations. Linkage conservation is also not as pervasive as the conservation of gene content, and a few eutherian lineages do exhibit some rearrangements (Amar et al. 1988; Piumi et al. 1998; Robinson et al. 1998; Sandstedt and Tucker 2004; Park et al. 2013). Conservation of gene order could simply be an indirect result of selective pressures to maintain gene content, or, conversely, intrachromosomal rearrangements could be directly selected against as they might disrupt some critical biological process like X chromosome inactivation.

Recent observations have also revealed striking similarities in regional rates of recombination along collinear X chromosomes between distantly related species (Li et al. 2018). Specifically, the X chromosome of cats, dogs, pigs and humans all share a large recombination cold spot that spans roughly one-third of the chromosome (Nagaraja et al. 1997; Wong et al. 2010; Ai et al. 2015; Li et al. 2016; Li et al. 2018), spanning from approximately 55 Mb to 118 Mb in the human X chromosome (Nagaraja et al. 1997; Li et al. 2018). Interestingly, this region is associated with strong selective sweeps and high levels of genetic differentiation in a number of mammalian taxa (Dutheil et al. 2015; Nam et al. 2015; Li et al. 2018; Lucotte et al. 2018). Like linkage conservation, it is still unclear what physical or functional constraints drive reduced recombination over such a large swath of the X chromosome in a majority of eutherian mammals studied.

Multispecies comparative approaches have the power to identify selective pressures that shape these unique aspects of X chromosome evolution. They are limited, however, by the fragmented nature of sex chromosome assemblies present in most draft genomes – a consequence of the highly repetitive, large ampliconic regions (multi-

kilobase segmental duplications with >99% pairwise identity) that cannot be accurately assembled using standard second-generation (and even third-generation) sequencing methods (Eichler et al. 2004; Alkan et al. 2011; Huddleston et al. 2014; Chaisson et al. 2015; Khost et al. 2017).

Mueller et al. (2013) conducted the first fine-scale comparative X chromosome comparison using finished-quality sex chromosome assemblies from the human and mouse to systematically test Ohno's law. They found that while most protein coding genes had orthologs in the other species (82% and 77% in human and mouse, respectively), there appeared to be rapid turnover of a subset of loci, most notably those found in the ampliconic regions. Apparent deviations from Ohno's law can be due to the loss of ancestral genes from the X chromosome, or from the acquisition of novel loci or lineage-specific duplications. Unfortunately, other mammalian X-chromosome assemblies that were available for comparison were highly fragmented, and the authors present the results with the caveat that classification of genes as lineage-specific or independently acquired could be biased by limited phylogenetic sampling. These biases could arise from such lineage-specific characteristics as mutation rates, effective population sizes, and mating systems/strategies, or lineage-specific historical events including large-scale structural rearrangements or the invasion of meiotic drivers or novel selfish genetic elements. In order to ensure that any inferences made about general patterns within X-chromosome evolution are not skewed by such biases, we should expand the phylogenetic breadth of sampling when conducting additional comparative studies.

Here we sought to expand the repertoire of available high-quality X-chromosome assemblies by producing a near complete assembly of the domestic cat X chromosome. Subsequently, we used this assembly in conjunction with other available high-quality X chromosome assemblies to conduct comparative analyses using species that span the mammalian phylogeny in order to address three fundamental questions. 1) Given the highly rearranged nature of the mouse X chromosome, as well as the greater degree of linkage conservation shared between humans and more distantly related species, is the high-turnover in ampliconic gene repertoire a common feature? Or, did the intrachromosomal rearrangements observed in the mouse X chromosome permit or facilitate a shift in the ampliconic gene catalog? 2) What are the possible physical or functional properties or constraints that have led to the unparalleled linkage conservation seen on the mammalian X chromosome? 3) Why do such distant lineages share a similar recombination cold spot? Does this correlate with the conservation of gene order and/or physical constraints imposed by processes like X chromosome inactivation (XCI) or meiotic sex chromosome inactivation (MSCI)?

3.2 Results

3.2.1 BAC clone sequencing, assembly, and annotation

We sequenced and assembled 83 X-linked BAC clones selected from regions containing gaps in the domestic cat FelCat8.0 X chromosome assembly. The assembled clones were mapped to the recently completed FelCat9.0 assembly, which was constructed using Pacific Biosciences (PacBio) SMRT sequencing technology. As the v9.0 assembly utilized long-read data, the number of gaps was reduced by two orders of

magnitude, from 5,592 in FelCat8.0 to 54 in FelCat9.0 (Table 3.1). As such, many of the clones that were sequenced were mapped to non-gapped locations along the X chromosome, providing a means to confirm the accuracy of our clone-based assembly approach and the existing FelCat9.0 (Table B3.1). We were able to span six of the 54 gaps in FelCat9.0 assembly, including two ampliconic regions incorporating ~467 kb of ampliconic BAC-clone sequence data. In addition, we improved and increased the sequence content around twelve additional gaps, including eight within ampliconic regions, adding ~980kb of additional sequence data (prior to short-read correction of the entire X chromosome assembly) (Tables B3.1-B3.2). Twenty-five of the remaining gaps within the assembly are within or adjacent to (<100kb away) the ampliconic regions we identified, and five are within regions that align to human X-chromosome sequence that is ampliconic. This could result in an underestimate of the amount of ampliconic sequence contained on the domestic cat X chromosome, but the estimated gap-sizes within these regions are estimated to be less than 600 kb, keeping the amount of domestic cat ampliconic sequence closer to that of human (~6.8 Mb) than the mouse (~20.6 Mb).

We found eight novel protein-coding loci (open reading frame > 400 nucleotides) within the newly incorporated sequences. Two of these loci occur within ampliconic regions, and one is adjacent to the human ampliconic sequence starting at ~135 Mb. The

Table 3.1. Comparison of FelCat8.0 and FelCat9.0 X chromosome assemblies, as well as our improved 9.1 X chromosome assembly.

Feature	FelCat8.0	FelCat9.0	FelCat9.1
X-chromosome (ungapped) length	127,282,370 bp	130,557,009 bp	130,203,636 bp
X-chromosome gap length	4,403,103 bp	2,806,785 bp	2,593,726 bp
X-linked unplaced contigs/scaffolds	6,350	26	17
No. gaps in X-chromosome assembly	5,592	54	47

cat shares considerable overlap in gene content with this human amplicon (Table B3.3), but contains no ampliconic sequence itself. The remaining five loci were found at the end of the domestic cat X chromosome (~129Mb), originating from the BAC clone FCAB-331E2, again adjacent to a region that is ampliconic in the human but not found to be ampliconic within our study (Table B3.3).

Five novel loci had no significant matches to the NCBI non-redundant protein sequence database. The remaining three loci had significant matches, including one sharing sequence 98.21% similarity to an annotated olfactory receptor in the cheetah, and another that shares 91.0% identity (49% query coverage) to a cancer/testis antigen 1-like paralog in the mountain lion. The last transcript with a significant BLAST alignment shares 89.01% identity with a DUF1725 domain-containing protein from *Neisseria polysaccharea* and 69.66% identity with an unnamed protein product from the domestic dog. Five of the eight novel loci displayed the highest levels of expression within the testes (Table B3.4), including the cancer/testis antigen 1-like locus.

3.2.2. Intraspecific X chromosome comparisons

Pairwise nucleotide alignments of the X chromosomes of the domestic cat, pig, and mouse to the human X chromosome confirmed conservation of gene order between the human, domestic cat, and pig across the entire length of the chromosome (Fig. 3.1; Tables B3.3,B3.5), as well as the numerous rearrangements of the mouse X chromosome (Pevzner and Tesler 2003b; Pevzner and Tesler 2003a; Mueller et al. 2013) (Figs. A3.1-A3.3; Table B3.4). The order of the human, domestic cat, and pig X chromosomes have been retained from the ancestral eutherian X chromosome, which has also been retained in the horse and elephant (Raudsepp et al. 2004; Delgado et al. 2009). We identified the ampliconic regions of each X chromosome through self-alignments using the standard definition of these regions ($\geq$ 99% identity with a length $\geq$ 10 kb; Mueller et al. 2013). Using these strict parameters, we identified far less ampliconic sequence than previously reported for the human and mouse X chromosomes (Mueller et al. 2013; Lucotte et al. 2018): ~2.5 Mb of human ampliconic sequence, and ~12.8 Mb of mouse ampliconic sequence; whereas by comparison Mueller et al. (2013) reported 3.15 and 19.4 Mb of ampliconic sequence, respectively. However, these previous studies included intervening sequence between ampliconic stretches as long as that distance was <500 kb in length. When we calculated total ampliconic sequence using this same definition, the total amount of ampliconic sequence within the human X chromosome rose to ~6.8 Mb and the mouse to ~20.6 Mb (Tables B3.6-B3.7). These estimates are based on newer genome assemblies than used in previous studies, which likely accounts for the increased

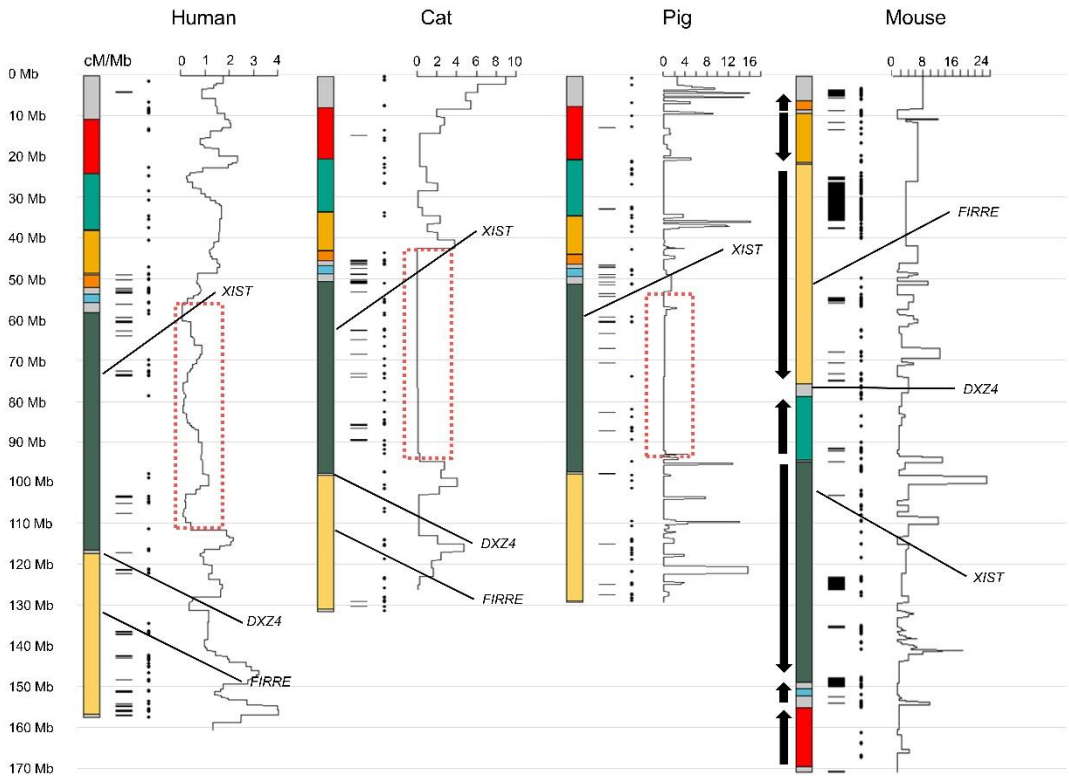


Figure 3.1. Interspecific comparison of four mammalian X chromosomes. Ideograms of the human, cat, pig, and mouse X chromosomes are depicted, with areas of conserved synteny shown in colored bands. For each X chromosome, the ampliconic regions are shown to the right as black bars, and the locations of lineage-specific gene gains are to the right of these bars depicted as black dots. Regional rates of recombination are plotted along the length of each X chromosome, and the shared recombination desert is outlined in the dashed red box for the human, domestic cat, and pig. The locations of loci prominently involved in X chromosome inactivation are given for the human, cat, and mouse, but are not yet annotated for the pig.

ampliconic content found in our study for the human and mouse X chromosomes. In order to facilitate comparisons to these previous studies, we applied the latter, less conservative parameters when delineating the boundaries of the ampliconic regions within the domestic cat and pig X chromosomes.

The domestic cat X chromosome assembly contains ~3.6 Mb of ampliconic sequence, with a majority of these regions located within the middle third of the chromosome (Fig.3.1; Table B3.8). The pig X chromosome displays the smallest amount of ampliconic sequence with only ~1.9 Mb, and, as in the domestic cat, these regions were concentrated within the middle third of the chromosome (Fig. 3.1; Table B3.9). The size and distribution of ampliconic regions appear similar between the human, cat, and pig X chromosomes, while the ampliconic regions of the mouse X chromosome are much larger and more widely dispersed.

X-linked protein-coding gene number varied between the four species, ranging from 774 in the pig to 943 in the mouse. The mouse X chromosome was an outlier in the total number of genes, amount of ampliconic sequence content, and in the number of lineage-specific gene gains and lineage-specific gene losses (Fig. 3.2; Table 3.2). The remaining three species were far more similar to one another showing both fewer putative gene gains, fewer instances of lineage-specific gene loss, and less total ampliconic sequence content.

Mueller et al. (2013) showed that lineage-specific gene gain events were enriched for ampliconic loci. We observed the same pattern in the human ($p < 0.0001$, $\chi2=69.49$; Table B3.10), mouse ($p < 0.0001$, $\chi2=59.87$; Table B3.11), the cat ($p =0.04393$ $\chi2=4.0592$; Table 3B.12), and pig ($p = 0.016$, $\chi2=5.797$; Table B3.13). The position of the ampliconic regions frequently appeared to be conserved across lineages.

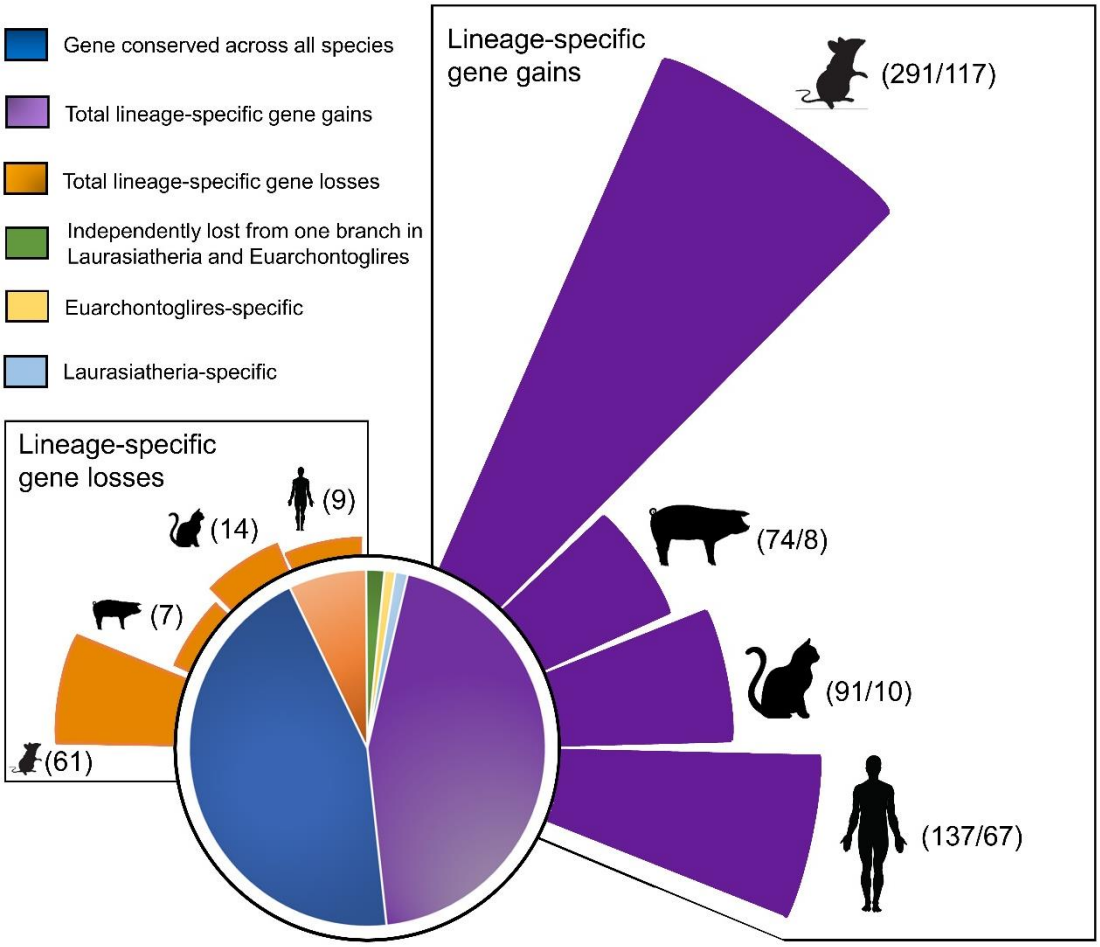


Figure 3.2. X-linked genes identified in this study across four mammal species (*n*=1,331) and their histories across lineages. Within the lineage specific gains and losses, the number of genes within these categories for each species is shown in parentheses. Within the lineage-specific gains, the number of these genes that are ampliconic are shown to the right of the total number of genes within the parentheses.

For example, the ampliconic region occurring between 51,668,020 bp and 52,920,401 bp of the human X chromosome shares genes with both the domestic cat and the pig,

Table 3.2. Content summary of the human, domestic cat, pig, and mouse X chromosomes.

Feature	Human	Cat	Pig	Mouse
X chromosome length (Mb)	150,040,895	130,557,009	125,939,595	171,030,299
X ampliconic length - 500kb inclusive (Mb)	6,804,303	3,601,968	1,884,848	20,608,963
X ampliconic length - non-inclusive (Mb)	2,499,509	1,039,382	712,937	12,760,152
Ampliconic gene number	136	30	16	123
Novel gene number	137	91	74	299
Novel ampliconic gene number	67	10	8	117
Novel single copy gene number	8	33	16	37
Novel multicopy gene number	62	48	50	139

although each lineage does appear to have acquired novel genes within the orthologous region. Furthermore, genes shared between species fall within ampliconic regions in some species, but fall just outside of the ampliconic boundaries in others (Fig. 3.3a). Similarly, an ampliconic region near the distal end of the human X chromosome (~154 Mb), contains genes conserved across all four species, some of which fall within an ampliconic region of the human X chromosome, but outside of the ampliconic regions of the domestic cat and mouse (Fig. 3.3b; Table B3.3).

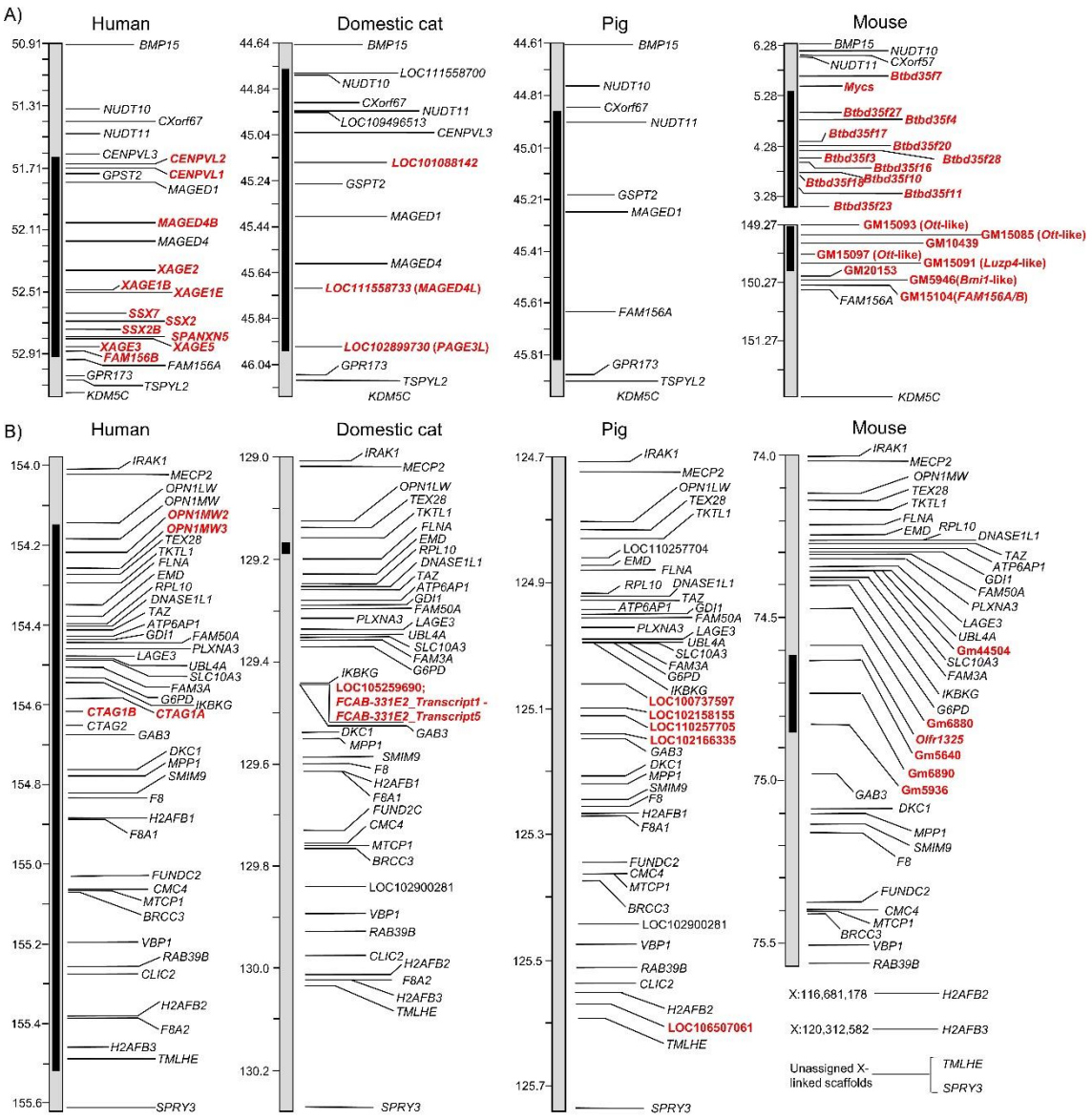


Figure 3.3. Annotation alignments of regions of the human, cat, pig, and mouse X chromosomes. A) Alignment of X chromosome regions corresponding to the human X chromosome around 51Mb-53 Mb. The mouse X chromosome has been rearranged within this region and the two disparate regions are shown. B) Alignment of X chromosomes corresponding to 154-155.6 Mb. Genes located within the distal portion of the alignment are found at non-contiguous regions within the mouse X chromosome and their positions within the mouse assembly are depicted separately. Ampliconic regions are delineated by black bars within the ideograms and lineage specific loci are shown in red.

Table 3.3. Ampliconic regions of the human, domestic cat, pig, and mouse X chromosomes. Ancestral regions (regions aligning within <1Mb of an ampliconic region with another species and/or containing the same genes) are assigned a number (ID) corresponding with Figure 3.3. Ampliconic regions which are lineage-specific are denoted with a dash (-).

Human Ampliconic Regions				Domestic cat Ampliconic Regions			
Start	End	Length	ID	Start	End	Length	ID
3,821,643	3,938,401	116,758	-	14,366,026	14,383,190	17,164	-
48,366,973	48,428,451	61,478	13	44,753,753	46,069,961	1,316,208	1
49,528,395	49,605,060	76,665	-	46,815,413	46,832,194	16,781	16
51,668,020	52,920,401	1,252,381	1	48,269,298	48,445,490	176,192	2
55,453,508	55,480,140	26,632	2	49,525,430	50,521,076	995,646	3
58,716,337	58,846,090	129,753	3	52,521,561	52,531,618	10,057	4
59,467,608	59,975,932	508,324	-	61,812,206	61,977,473	165,267	5
61,974,100	61,984,719	10,619	-	64,216,861	64,230,511	13,650	17
63,129,000	63,224,416	95,416	4	67,748,352	67,765,210	16,858	18
71,683,102	71,798,461	115,359	5	72,528,457	72,564,863	36,406	19
72,740,931	72,979,901	238,970	5	73,282,593	73,353,509	70,916	19
102,197,558	102,489,558	292,000	6	84,809,503	85,235,500	425,997	6
103,953,836	103,989,433	35,597	6	85,758,896	85,813,462	54,566	6
106,266,280	106,278,816	12,536	7	88,576,818	88,840,446	263,628	7
115,843,208	115,886,139	42,931	8	128,107,195	128,118,411	11,216	11
120,038,203	120,198,162	159,959	14	129,175,617	129,187,033	11,416	12
120,929,382	120,987,033	57,651	-				
135,116,125	135,889,734	773,609	9				
141,002,115	141,585,255	583,140	-				
146,801,758	146,822,756	20,998	15				
149,573,132	149,796,256	223,124	10				
152,678,581	153,293,926	615,345	11				
154,149,491	155,504,549	1,355,058	12				

Table 3.3. Continued.

Pig Ampliconic Regions				Mouse Ampliconic Regions			
Start	End	Length	ID	Start	End	Length	ID
12,227,750	12,314,340	86,590	-	3,321,011	5,313,997	1,992,986	1
31,653,173	31,869,295	216,122	-	8,480,304	8,499,226	18,922	13
45,054,183	45,858,289	804,106	1	11,298,326	11,328,120	29,794	-
47,435,127	47,474,791	39,664	16	13,167,131	13,197,566	30,435	-
48,022,185	48,041,616	19,431	2	24,808,958	25,346,966	538,008	-
49,181,699	49,206,013	24,314	3	25,994,206	35,451,485	9,457,279	-
49,881,516	49,894,192	12,676	3	37,251,522	37,402,518	150,996	14
51,972,371	52,070,728	98,357	-	54,272,650	55,791,158	1,518,508	9
52,611,043	52,622,268	11,225	-	67,680,296	67,694,529	14,233	15
57,640,099	57,657,993	17,894	5	70,430,244	70,509,140	78,896	10
58,656,069	58,832,442	176,373	5	73,118,084	73,148,456	30,372	11
61,474,449	61,484,687	10,238	17	74,642,652	74,840,177	197,525	12
65,080,691	65,191,728	111,037	18	91,447,322	91,965,791	518,469	-
68,586,834	68,597,679	10,845	19	94,706,195	94,816,711	110,516	-
80,370,632	80,383,729	13,097	-	103,003,108	103,016,033	12,925	5
84,779,597	84,806,994	27,397	6	123,113,661	126,051,901	2,938,240	-
95,013,653	95,139,430	125,777	8	135,107,360	135,564,906	457,546	6
111,987,560	112,007,873	20,313	9	147,824,892	150,099,691	2,274,799	8
121,615,135	121,648,641	33,506	10	152,409,657	152,438,240	28,583	1
124,023,929	124,049,815	25,886	11	154,000,736	154,066,465	65,729	-
				170,737,096	170,881,298	144,202	-

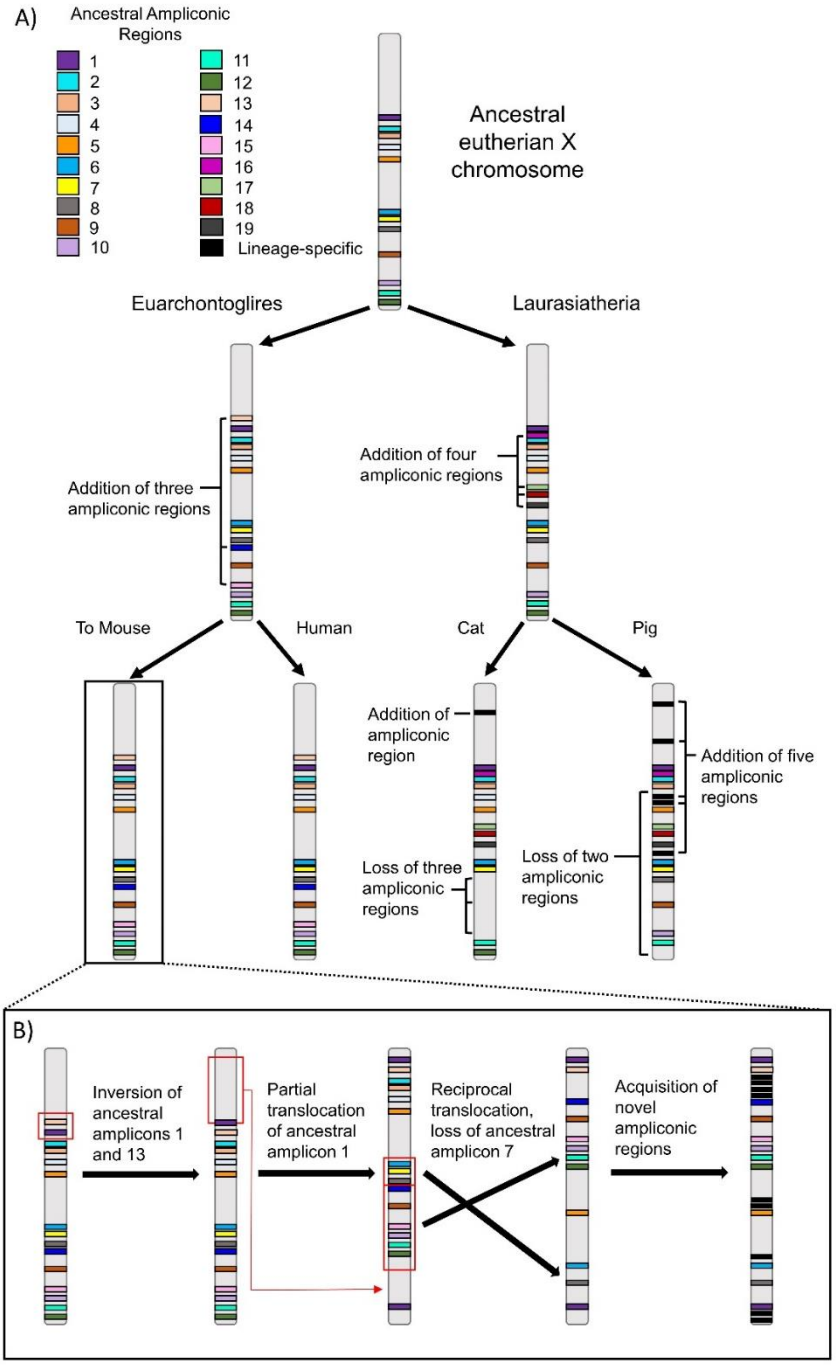


Figure 3.4. A) The evolution of ampliconic regions from the ancestral eutherian X chromosome to the human, pig, and cat X chromosomes. The addition of ampliconic regions after the Laurasiatheria-Euarchontoglires split could also be independent losses of these regions within the reciprocal group, and will require further sampling. The numbers assigned to ancestral ampliconic regions match those in Table 3.2. B) Simplified evolutionary path of the mouse X chromosome following split from human. Ampliconic regions are not drawn to scale.

We classified ancestral ampliconic regions based on two criteria: 1) conserved gene content and 2) physical proximity to ampliconic sequence found within at least one other species (i.e., within 1 Mb of aligned conserved genes between species). Using these criteria, we identified 17 of 23 ampliconic regions within the human as ancestral, 16 of 17 as ancestral within the domestic cat, 15 of 20 within the pig, and 12 ampliconic regions out of 21 as ancestral within the mouse (Fig. 3.4; Table 3.3). Mouse lineage-specific evolutionary breakpoints were frequently associated with ampliconic regions, often flanking or spanning the breakpoint in gene order (Fig. 3.3, 3.4; Table B3.3). For example, the ampliconic regions at ~ 3-5 Mb and ~149 Mb of the mouse, coincide with an ancestral eutherian ampliconic region located at ~51 Mb in human, which is also shared with the cat and pig (Figs. 3.3 and 3.4; Table 3.3 and B3.3). Interestingly, the other end of the mouse lineage inversion corresponds with yet another ampliconic region shared between the human and pig X chromosomes (Fig. 3.3; Table 3.3).

We used previously published data for the human, cat, pig, and mouse to map recombination rates along the lengths of each X chromosome (Kong et al. 2002; Ma et al. 2010; Li et al. 2016; Simecek et al. 2017) (Fig. 1). Coordinates for a previously identified recombination cold spot within the human, cat, and pig (Nagaraja et al. 1997; Fernández et al. 2014; Li et al. 2016; Li et al. 2018) were enriched for ampliconic regions within the cat ($p = 0.0034$, $\chi^2 = 8.533$), but not within the human ($p = 0.76$, $\chi^2 = 0.093$) or the pig ($p = 0.744$, $\chi 2 = 0.107$).

3.3 Discussion

3.3.1. Ampliconic regions promote novel gene acquisition across lineages

In 1967, Susumu Ohno first proposed what has now become known as Ohno's law: the gene content of the mammalian X chromosome should remain highly conserved owing to the need to maintain relative levels of expression between X-linked genes and autosomal loci. Mueller et al. (2013) published the first fine-scale test of this law by comparing the gene content of two high quality mammalian X chromosome reference assemblies from human and mouse, both members of the superordinal lineage Euarchontoglires (Murphy et al. 2001). They revealed an apparent dichotomy in the evolution of the X chromosome: a majority of the single-copy X chromosome gene content was conserved between these two species, but a majority of ampliconic genes were lineage specific and functioned primarily within the male germ line. However, these comparisons were phylogenetically restricted to just two members of the superordinal lineage Euarchontoglires, and this analysis included mouse, a species with the most accelerated rates of nucleotide substitution and genome evolution within eutherian mammals (Springer and Murphy, 2007).

Here we expanded this phylogenetic scope to add representatives of two additional, divergent eutherian orders, Carnivora (cat) and Cetartiodactyla (pig), both from the superordinal clade Laurasiatheria, and with X chromosomes that are strongly collinear with human and the putative eutherian ancestor. While our findings do agree with Mueller's observation that the ampliconic regions of the X chromosome are enriched for lineage-specific gene gains, we point out several exceptions and

refinements to their hypothesis of X chromosome evolution and specialization. First, the acquisition of novel or lineage-specific genes is not an exception or violation of Ohno's law, as species can conserve the original gene content of the ancestral X chromosome as well as acquire new gene content via translocation, *de novo* evolution, or expansion of ancestral gene family members. Furthermore, as a significant majority of acquired genes are expressed predominantly within the male germ line (Mueller et al. 2013), they would not have an impact on dosage compensation within females.

Second, our results suggest that the evolution of eutherian ampliconic regions has not been dominated by independent acquisition, as these regions frequently appear to occur in positional orthology across the length of the chromosome, and often contain either the same genes or paralogous members of the same gene family. Thus, our results suggest that a majority of the ampliconic regions found in eutherian X chromosomes were likely present in the eutherian ancestor, and did not evolve following the KPg radiation of eutherian orders (Mueller et al. 2013). Ampliconic regions are more prone to rapid turnover than are the single copy regions of the X chromosome, and may possess some physical or structural characteristics that make them more likely to acquire additional ampliconic sequence. For example, the boundaries of ancestral topologically associating domains (TADs) are highly conserved across mammals and serve as a structural mechanism by which genes are regulated; as such, disruption of TAD boundaries can cause changes in gene regulation that have drastic phenotypic effects (Dixon et al. 2012; Nora et al. 2012; Lupiáñez et al. 2015; Galupa and Heard 2017; Krefting et al. 2018). As repetitive regions are prone to acquire large structural

variations, including deletions or inversions (Richards et al. 2005; Hastings et al. 2009; Sudmant et al. 2015), and as such changes in structure can have dramatic deleterious consequences if they disrupt TADs, the acquisition of ampliconic sequence outside of these conserved domains may be more readily permitted than within them. Comparative studies comparing TAD boundaries and the distribution of ampliconic sequence within the X chromosome are necessary to evaluate this hypothesis.

3.3.2. Extensive conservation of gene order across distantly related taxa

Ohno's law predicted the conservation of gene content of the ancestral mammalian X chromosome across the clade, but does not necessarily invoke the conservation of gene order. While the degree of conservation of X-chromosome gene order is less than that of gene repertoire, the extent of conserved synteny between mammalian X chromosomes is far greater than seen between any mammalian autosomes (Murphy et al. 1999; Murphy et al. 2005; Delgado et al. 2009). While collinearity may simply be an indirect consequence of the constraints on gene content during the early evolution of dosage compensation, the large divergence times separating the majority of species that display X-chromosome collinearity suggest that other selective pressures are likely contributing to this stability.

Previous studies have suggested that the physical constraints imposed by the mechanics of XCI could be the driving factor behind the extensive collinearity observed between mammalian X chromosomes (Delgado et al. 2009). X-chromosome inactivation in eutherian mammals involves the spread of *XIST* transcript, both linearly out from the

X-chromosome inactivation center, and through three-dimensional interactions within the X chromosome being silenced (Plath et al. 2002; Engreitz et al. 2013). This process could be hindered by rearrangement of X-chromosome gene order.

Our results show that the collinearity of X-linked loci is not restricted to single copy genes, but frequently extends to ampliconic regions as well (Fig 3.3; Table B3.3). Furthermore, the amount of X-linked ampliconic sequence within the human, domestic cat, and pig was far less than that observed within the mouse. This may suggest evolutionary constraints on the amount of ampliconic sequence that may be acquired on the X chromosome, in addition to whatever selective pressures appear to constrain the spacing and distribution of these regions along the X chromosome. If true, then the rearrangements that occurred on the ancestral branch leading to the mouse X chromosome may have facilitated or permitted the expansion of ampliconic regions within this lineage. As sperm competition is likely an important component of sexual selection within *Mus* (Dean et al. 2006), and a significant portion of novel ampliconic loci have evolved for the male germ line (Mueller et al. 2008; Mueller et al. 2013), any release from constraint on the amount of X-linked ampliconic sequence may have allowed for the mouse to drastically expand gene families specialized for the male germ line.

The massive expansion of ampliconic sequence within the mouse X chromosome could also have been compensatory, as such large scale rearrangements would likely have deleterious effects. Some of the larger ampliconic regions identified within the mouse X chromosome occur within the blocks of conserved linkage (Fig. 3.1). These

novel expansions may facilitate interactions within the mouse X chromosome that aid in the process of XCI, or promote proper spacing between loci involved in this process. The relative positions of several loci involved in XCI are dependent on their order and orientation, namely *DXZ4*, *XIST*, and *FIRRE* (Bonora et al. 2018). These loci are retained within the same order in the mouse as in the human and domestic cat (Fig. 3.1). What selective pressures exist to maintain this spatial orientation and the history of rearrangements of the mouse X chromosome is still unclear, but our results suggest that the ancestral eutherian ampliconic regions have acted as hot spots for evolutionary breakpoints within the mouse.

3.3.3. Ampliconic regions and linkage conservation

Ampliconic regions are enriched at the evolutionary break points observed between the derived arrangement of the mouse X chromosome, and the conserved X chromosomes of the human, cat, and pig (Fig 3.1, 3.3; Table 3.2; Table B3.3). If the ampliconic regions are only found between or immediately adjacent to evolutionary breakpoints in the derived state, then the most parsimonious explanation of their presence would be their derivation after rearrangement. Conversely, if the ampliconic regions are also present at orthologous regions in species maintaining the ancestral gene order, then it is conceivable that amplicons were present in the ancestral state. We identified several instances in which the ampliconic regions were neighboring or spanning breaks in synteny, not only in the mouse but also in the other species examined (Fig 3.1.; Table B3.3). This suggests that these regions act as breakpoint hotspots along

the X chromosome at which rearrangements are most likely to occur. The correlation between repetitive regions and evolutionary breakpoints has been documented in other species. For example, segmental duplications are enriched around synteny breakpoints between the human and gibbon (Carbone et al. 2014). To confirm the patterns we observed in X chromosome breakpoints within this study, this would have to be examined across a larger number of taxa with derived X chromosome organizations.

3.3.4. Conclusions

By sampling taxa across a broader evolutionary time-scale, we have shown that the majority of ampliconic regions of mammalian X chromosomes were ancestral, and did not originate de novo as part of lineage-specific innovations. While amplicons do appear to undergo more rapid rates of evolution than the conserved single-copy regions of the X chromosome, they are not necessarily evolving independently across all mammalian lineages as 59 of 80 ampliconic regions were ancestral. Our results suggest that the mouse is likely a poor model for understanding X chromosome evolution in general, as it is a clear outlier in not only gene order, but in the acquisition of numerous lineage-specific genes and the loss of ancestral alleles. We posit that these two main differences, extensive gene gain/loss and structural rearrangements, are likely closely intertwined, and that the mouse may have either exploited a lack of constraint on amplicon size after structural rearrangements by massively expanding the X-linked gene repertoire involved in male gametogenesis, or had to undergo rapid compensatory

adaptation to recover from breaking the physical constraints imposed by X chromosome inactivation. If the latter is true, then the use of ancestral ampliconic regions as evolutionary breakpoints to restore the orientation and proximity of loci involved in XCI may have contributed to the pattern of X chromosome reorganization within the mouse. Future studies including species with derived X chromosomes (e.g., cow, deer) could elucidate whether this pattern holds and may shed more insight into the physical constraints governing eutherian X chromosome evolution.

3.4 Methods

3.4.1 BAC clone sequencing, assembly, and annotation

We conducted sequencing and assembly of BAC clones from the *Felis catus* female BAC library FSCC from Amplicon Express. Clones were chosen based on the mapping location of the BAC end sequences aligned to the domestic cat felCat8.0 genome assembly. Briefly, clones with both end sequences mapping to either side of a gap within the assembly, or a single end mapped adjacent to a gap, were selected for sequencing. Selected clones were cultured and DNA was extracted using standard protocols. Clone DNA was pooled into three separate groups and sequenced using the PacBio Sequel system.

Given the previously described disparities in the abilities of different assembly pipelines and parameters to accurately reconstruct complex regions of the genome (Khost et al. 2017), we assembled each pool separately using both the Celera 8.3rc2 (Myers et al. 2000; Koren et al. 2012; Koren et al. 2013; Berlin et al. 2015) and Canu

(Koren et al. 2017) pipelines with a variety of parameters. Raw PacBio reads were mapped to each assembly using the program BLASR with default settings (Altschul et al. 1990), and the resulting alignment was used to refine each assembly using Arrow within the GenomicConsensus software from PacBio.

In order to remove any sequences not originating from the domestic cat genome, each assembly was aligned to the E. coli genome using BLAST (Altschul et al. 1990) in Geneious (Kearse et al. 2012), and the resulting alignments were examined by eye to confirm and remove any contaminants. The vector sequence used in the BAC library was then mapped to the remaining sequences using LASTZ (version 1.02.00, Harris 2007), and a custom script was used to remove any vector present in the assembled sequences.

Next, we downloaded all available BAC-end sequences from NCBI and mapped these to each assembly, allowing us to identify ends of separate BAC clones that had been assembled together due their overlap within the genome. These assemblies were then aligned using the program MAFFT (Katoh and Standley 2013) and were visually inspected to identify any major disparities. If none were found, the consensus sequence from the alignment was used as the representative sequence for the clone. In those cases where the different assembly pipelines produced assemblies with major disparities, the consensus sequence for the longer sequences was used, as misassembly of ampliconic regions usually results in the collapse of adjacent segmental duplications.

The assembly for each clone was then incorporated into the X-chromosome scaffold of the PacBio long read genome assembly (version 9.0) based on the mapping

of BAC-end sequences and BLAST alignments. We removed any unincorporated contigs from the genome assembly file that appeared to be covered by our sequenced clones, and mapped Illumina whole-genome sequence data from SRR5055389 onto the new genome that included incorporated clone assemblies using the program BWA-MEM (Li and Durbin 2009). We then used the resulting alignment files and the program Pilon (Walker et al. 2014) to correct the assembly. The quality of the incorporated BAC clone sequences were then assessed by mapping Illumina whole genome sequence data from the domestic cat to the updated assembly and subsequently checking mapping statistics using the program SAMtools (Li et al. 2009) (Supplemental Table S3).

After we completed incorporating the assembled BAC clone sequences, we aligned RNA-seq data from two testis (SRR1981105 and SRR3200462), cerebellum (SRR3218718) kidney (SRR3200460), heart (SRR3200471), lung (SRR3200449), and uterus (SRR3200458) tissues from the domestic cat using the program STAR (Dobin et al. 2013) using default parameters. Transcripts were subsequently constructed using the program Cufflinks (Trapnell et al. 2012). Assembled transcripts originating from the newly incorporated sequences were then assessed for protein-coding potential, requiring a minimum open reading frame of 400 nucleotides and detectable expression in at least two samples. These stringent parameters were chosen to minimize the possibility of falsely inflating the number of novel protein-coding genes within ampliconic genes, as the regions of the FelCat9.0 X chromosome assembly that were improved in this study consisted primarily of ampliconic content. The resulting transcripts were then aligned to the FelCat9.0 genome using BLAST (Altschul et al. 1990) to ensure they were not

already present within the FelCat9.0 annotation, as well as to assign chromosome coordinates to the loci for incorporation into our multi-species alignments of X chromosome annotations.

3.4.2. Intraspecific X chromosome comparisons

We downloaded the annotation files for the human (GRCh38.p12), domestic cat (Felis_catus_9.0), pig (Sscrofa11.1) and mouse (GRCm38.p6) genomes, and manually aligned the X-linked annotations (Supplemental Table S4) in Excel. In order to confirm the results of our manual alignment of the annotation files, we conducted pairwise alignments of the cat, pig, and mouse X chromosomes to the human X chromosome using the program LASTZ (Harris 2007). Novel domestic cat loci that we identified in the study were aligned to the FelCat9.0 X chromosome in order to identify their position within the annotation alignments. If the loci originated from X chromosome sequence that was not in the FelCat9.0 assembly, or that were located on unassigned scaffolds/contigs, the coordinates of the sequence replaced in our updated assembly were used to assign position.

We identified ampliconic regions by conducting self-alignments of each of the X chromosomes (human (GRCh38.p12), domestic cat (Felis_catus_9.0), pig (Sscrofa11.1) and mouse (GRCm38.p6)) using the program Nucmer (Delcher et al. 2002) using the "--maxmatch" and "--nosimplify" parameters. Alignments were then filtered to remove any self-aligned sequences, or sequences that were <99.0% or <10 kb in length. We then clustered ampliconic regions that were within 500 kb of one another, as in Mueller et al.

(2013), in order to facilitate comparisons between studies (Supplemental Tables S6-S9). Figures for the annotation alignments, including the locations of ampliconic loci, recombination rates, and novel genes were constructed in R statistical program using the karyoploteR package (Gel and Serra 2017).

Expected values for all Chi-square analyses were normalized to account for differences in the lengths of ampliconic and non-ampliconic regions, or between the length of the recombination desert and the two flanking regions. For example, when testing for the enrichment of novel loci in ampliconic regions, if the X chromosome was comprised of 5% ampliconic sequence and contained 100 ampliconic genes, the expected number of novel genes to be ampliconic was 5 and non-ampliconic 95.

CHAPTER IV

CONCLUSIONS AND FUTURE WORK

The main impetus behind the research described within this dissertation was to expand our knowledge of mammalian sex chromosome evolution beyond human and mouse, and, as such, provide a better basis on which to examine the evolution of reproductive isolation. As it has been over a decade since the first (and last) mammalian hybrid sterility locus was fine-mapped to a definitive gene (Mihola et al. 2009), it is our hope that the resources generated through this dissertation research will aid in the discovery of additional genes contributing to reproductive isolation and speciation in mammals. The expansion of our knowledge regarding sex-chromosome evolution will help inform future studies into the origin of reproductive isolation within the felid system, as the X chromosome is known to play a disproportionate role in this phenomenon (Large X-Effect). Furthermore, the results we obtained by conducting comparative analyses, both within Felidae and across superordinal clades of mammals, have refined and challenged the conclusions of earlier research based on the study of one or two species. These earlier studies have been used to infer large-scale evolutionary processes that do not hold across the breadth of Mammalia (>6,000 species). We have shown that the evolution of sex chromosomes within this class is far more complicated than previously appreciated, and that incorporating a broader phylogenetic sampling approach is imperative before any overarching patterns of sex-chromosome evolution can be confidently inferred.

4.1 The Y is dead, long live the Y

The evolutionary trajectory of the Y chromosome has been the source of contentious debate for decades (Graves 2016; Abbott et al. 2017). Once thought to be nothing more than a metaphorical ghost ship of junk DNA and male sex determination, advances within the last decade have shown that the mammalian Y chromosome is relatively stable, comprising a suite of ancestral, widely-expressed, dosage-sensitive genes (Bellott et al. 2014; Cortez et al. 2014). Although the Y underwent a period of rapid decay and loss of genes during the early stages of mammalian evolution, the rate of decay has all but come to a complete halt (Bachtrog et al. 2013).

Our study of Y chromosome evolution within Felidae demonstrated that Y chromosome stability and conservation can extend beyond the single copy region into ampliconic regions. Although these regions exhibit rapid turnover in genic content between some species, we show that the Y ampliconic gene repertoire can remain constant over tens of millions of years, even in the face of large-scale structural rearrangements and variations in gene family copy number. Furthermore, we show that these genes can be retained over long evolutionary timespans even as the Y chromosome recruits new gene families and regulatory elements from other parts of the genome (Chapter II).

4.2 Y chromosome evolution in South American felids

An interesting observation that arose from our study of felid Y chromosomes was the lack of FISH hybridization signal covering the distal portion of the tigrina Y

chromosome, using probes that covered the entire Y of members from all other felid lineages. These probes, which were derived from known felid ampliconic Y genes, hybridize across most, if not all, of the Yq of all other species included in these analyses. Lack of probe hybridization to the in tigrina suggests extensive acquisition of novel ampliconic gene content within the South American felid clade. While this is interesting in itself, it is noteworthy that several members of this clade exhibit signs of historic and ongoing hybridization, including the tigrina (*Leopardus tigrinus*), Geoffroy's cat (*L. geoffroyi*), and, the pampas cat (*L. colocolo*) (Trigo et al. 2008; Trigo et al. 2013; Trigo et al. 2014). The Y chromosome is known to contribute to sperm abnormalities (Repping et al 2002), but it has not been implicated in contributing to reproductive isolation (although it does contribute to sperm abnormalities in hybrid mice – Campbell and Nachman 2014). This system therefore could provide insight into the structural and functional properties of recently evolved Y-linked ampliconic sequence within a hybridizing background.

4.3 A putative meiotic drive element within Panthera

We documented the co-expansion of the Y-linked ampliconic gene *CCDC71LY* and its autosomal progenitor *CCDC71L* within the *Panthera* (Chapter II). We discussed the possibility of these loci acting together in some reproductive function as *CCDC71L* was highly expressed within the fallopian tube as well as in the testes, and *CCDC71LY* displayed testis-specific expression. We also discussed the possibility that these two loci could be interacting members of a meiotic drive/repression system as seen between *Sly*

and *Slx* in *Mus*, although within *Panthera* one member of the meiotic drive pair would be autosomal.

The putative role of *CCDC71L* and *CCDC71LY* was supported by the presence of sex-biased ratios in the litters of hybridizing lions and tigers (Table B2.1) However, the presence of skewed sex-ratios in the opposite direction have also been reported, albeit anecdotal. We also noted disparities in the copy number estimates of both *CCDC71LY* and *CCDC71L* in both the tiger and lion between our *in-silico* and qPCR based approaches. While this could simply be the result of disparities in sensitivity between the two approaches, the individuals used in each analysis were different. Given this, and the anecdotal reports of skews sex-ratios in hybrid litters in opposite directions, the *CCDC71LY*/*CCDC71L* pair could represent the presence of a CNV meiotic drive/suppression system within *Panthera.* Whether the co-expansion of these two gene families is involved in a cooperative interaction or indicates the presence of a meiotic drive system, the observation warrants further investigation.

4.4 The complex history of ampliconic regions is underappreciated

Our initial delineation of ampliconic regions on human and mouse X chromosomes identified far less ampliconic sequence than reported in previous studies (Mueller et al. 2013; Lucotte et al. 2018). These studies clustered ampliconic regions into single units if they occurred within 500 kb of one another. After noticing the drastic differences between our results and those of Mueller et al. and Lucotte et al. we adopted the same, far more permissive criterion in order to facilitate comparisons. While the

clustering of some truly ampliconic regions (>10kb in length, >99.0% shared identity; Mueller et al. 2013) makes sense, clustering regions as far away as 400 or 500 kb dramatically inflates estimates of the amount of "ampliconic" sequence present. This degree of inclusivity also leads to the classification of genes that are clearly ancestral single-copy genes as ampliconic. Furthermore, this also simplifies the history of ampliconic regions and creates a false dichotomy in sequence classification and fails to appreciate the complexities observed throughout the evolution of these regions.

For example, the largest ampliconic region within the domestic cat X chromosome was found between base pairs 44,753753 and 46,069,961, spanning a total length of over 1.3 Mb. Within this region, however, approximately 220kb of nucleotide sequence meet the strict definition of ampliconic (>10kb in length, >99.0% shared identity; Mueller et al. 2013) (Table B3.7). While this region still contains four gaps that would likely cause us to underestimate the amount of ampliconic sequence, the combined lengths of the estimated gap sizes is less than 40kb, with one gap comprising a vast majority of that estimate (37,643 estimated bps).

The distribution and orientation of the ampliconic sequences within this region show that it is comprised of a series of distinct palindromes and segmental duplications. Furthermore, this region contains a number of distinct gene families that have been differentially expanded among lineages. In humans this region exhibits lineage-specific expansions of the *CENPVL* (centromere protein C like) family, as well as the *XAGE* and *SSX* gene families (Table B3.3). A cluster of *MAGE* genes have also been expanded within this region in both human and domestic cat. Further demonstrating the complexity

of amplicon evolution, both *NUDT10* and *NUDT11*, genes shared between all species examined in this dissertation, fall within ampliconic regions of the domestic cat, but not in the human, pig, or mouse.

This specific example highlights the complex history of the evolution of ampliconic regions, and demonstrates the need for additional study from a greater diversity of species in the future. It also demonstrates the need to be clear when discussing ampliconic regions. Although they are often clustered within regions, the smaller units that make up the inclusive ampliconic region have their own evolutionary trajectories. This is evidenced by the lineage-specific gene gains/losses in shared and unique ampliconic gene families.

4.5 Ampliconic regions act as genomic fault lines

We showed that X chromosome ampliconic regions are enriched for lineage-specific gene gains in the human, domestic cat, pig, and mouse (Chapter III). We also observed an enrichment of evolutionary breakpoints within or adjacent to ampliconic regions from the highly rearranged mouse X chromosome. These two findings suggest that the ampliconic regions of the mammalian X chromosome act as metaphorical genomic fault lines, where the evolution of lineage specific gene expansions alter the immediate genomic landscape and breaks in linkage and large-scale rearrangements change the overarching architecture of the chromosome.

It remains unclear, however, what physical or molecular properties promote the existence of these fault lines. Recent advances in sequencing technologies (e.g. Hi-C)

have shed light on higher order, three-dimensional organization of the genome (Bonev and Cavalli, 2016). Topologically associating domains (TADs) are one such level of genome organization, acting as functional units that constrain the co-regulation of gene networks (Dixon et al. 2016). Despite the increasing number of studies investigating TADs and other long-range interactions, there is still much to learn about their evolution and function (Bonev and Cavalli, 2016; Ramirez et al. 2018). During X-chromosome inactivation, the inactive X forms a bipartite structure that contains two massive superdomains (Deng et al. 2015). We propose that these and additional smaller domains likely contribute to the extensive conservation of gene order of the mammalian X chromosome, as well as to the distribution of ampliconic regions and the physical constraints which govern their evolution.

REFERENCES

Abbott JK, Nordén AK, Hansson B. 2017. Sex chromosome evolution: historical insights and future perspectives. *Proc R Soc B* **284**: 20162806.

Abyzov A, Urban AE, Snyder M, Gerstein M. 2011. CNVnator: An approach to discover, genotype, and characterize typical and atypical CNVs from family and population genome sequencing. *Genome Res* **21**: 974-984.

Ai H, Fang X, Yang B, Huang Z, Chen H, Mao L, Zhang F, Zhang L, Cui L, He W et al. 2015. Adaptation and possible ancient interspecies introgression in pigs identified by whole-genome sequencing. *Nat Genet* **47**: 217-225.

Alkan C, Sajjadian S, Eichler EE. 2011. Limitations of next-generation genome sequence assembly. *Nat Methods* **8**: 61-65.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *J Mol Biol* **215**: 403-410.

Amar LC, Dandolo L, Hanauer A, Cook AR, Arnaud D, Mandel JL, Avner P. 1988. Conservation and reorganization of loci on the mammalian X chromosome: a molecular framework for the identification of homologous subchromosomal regions in man and mouse. *Genomics* **2**: 220-230.

Augui S, Nora EP, Heard E. 2011. Regulation of X-chromosome inactivation by the X-inactivation centre. *Nat Rev Genet* **12**: 429-442.

Bachtrog D. 2006. A dynamic view of sex chromosome evolution. *Curr Opin Genet Dev* **16**: 578-585.

Bachtrog D. 2008. The temporal dynamics of processes underlying Y chromosome degeneration. *Genetics* **179**: 1513-1525.

Bachtrog D. 2013. Y-chromosome evolution: emerging insights into processes of Y-chromosome degeneration. *Nat Rev Genet* **14**: 113-124.

Band MR, Larson JH, Rebeiz M, Green CA, Heyen DW, Donovan J, Windish R, Steining C, Mahyuddin P, Womack JE, et al. 2000. An order comparative map of the cattle and human genome. *Genome Res* **10**: 1359-1368.

Barone MA, Roelke ME, Howard J, Brown JL, Anderson AE, Wildt DE. 1994. Reproductive Characteristics of Male Florida Panthers: Comparative Studies from Florida, Texas, Colorado, Latin America, and North American Zoos. *J Mammal* **75**: 150-162.

Barr ML, Bertram EG. 1949. A morphological distinction between neurons of the male and female, and the behaviour of the nucleolar satellite during accelerated nucleoprotein synthesis. In *Problems of Birth Defects*, (ed. Persaud TVN) pp. 101-102. Springer, Dordrecht.

Bateson W. 1922. Evolutionary faith and modern doubts. *Nature* **109**: 553-556.

Bellott DW, Cho TJ, Hughes JF, Skaletsky H, Page DC. 2018. Cost-effective high-throughput single-haplotype iterative mapping and sequencing for complex genomic structures. *Nat Protoc* **13**: 787-809.

Bellott DW, Hughes JF, Skaletsky H, Brown LG, Pyntikova T, Cho TJ, Koutseva N, Zaghlul S, Graves T, Rock S et al. 2014. Mammalian Y chromosomes retain widely expressed dosage-sensitive regulators. *Nature* **508**: 494-499.

Berlin K, Koren S, Chin CS, Drake JP, Landolin JM, Phillippy AM. 2015. Assembling large genomes with single-molecule sequencing and locality-sensitive hashing. *Nat Biotechnol* **33**: 623-630.

Bonev B, Cavalli G. 2016. Organization and function of the 3D genome. *Nat Rev Genet* **17**: 661-678.

Bonora G, Deng X, Fang H, Ramani V, Qiu R, Berletch JB, Filippova GN, Duan Z, Shendure J, Noble WS et al. 2018. Orientation-dependent Dxz4 contacts shape the 3D structure of the inactive X chromosome. *Nay Commun* **9**: 1445.

Brooks LD. 1988. The evolution of recombination rates. In Michod RE, Levin BR (eds) The Evolution of Sex (pp 87–105). Sinauer.

Bull JJ. 1983. Evolution of Sex Determining Mechanisms. Benjamin/Cummings.

Campbell P, Good JM, Dean MD, Tucker PK, Nachman MW. 2012. The contribution of the Y chromosome to hybrid male sterility in house mice. *Genetics* **191**: 1271-1281.

Campbell P, Nachman MW. 2014. X–Y interactions underlie sperm head abnormality in hybrid male house mice. *Genetics* **196**: 1231-1240.

Carbone L, Alan Harris R, Gnerre S, Veeramah KR, Lorente-Galdos B, Huddleston J, Meyer TJ, Herrero J, Roos C, Aken B et al. 2014. Gibbon genome and the fast karyotype evolution of small apes. *Nature* **513**: 195.

Carithers LJ, Ardlie K, Barcus M, Branton PA, Britton A, Buia SA, Compton CC, DeLuca DS, Peter-Demchok J, Gelfand ET et al. 2015. A Novel Approach to High-Quality

Postmortem Tissue Procurement: The GTEx Project. *Biopreserv Biobank* **13**: 311-319.

Carvalho AB, Clark AG. 2013. Efficient identification of Y chromosome sequences in the human and Drosophila genomes. *Genome Res* **23**: 1894-1907.

Chaisson MJP, Wilson RK, Eichler EE. 2015. Genetic variation and the de novo assembly of human genomes. *Nat Rev Genet* **16**: 627-640.

Charlesworth B. 1978. Model for evolution of Y chromosomes and dosage compensation. *Proc Natl Acad Sci U.S.A.* **75**: 5618-5622.

Charlesworth B. 1996. The evolution of chromosomal sex determination and dosage compensation. *Curr Biol* **6**: 149-162.

Charlesworth B, Charlesworth D. 2000. The degeneration of Y chromosomes. *Phil Trans R Soc Lond B* **355**: 1563-1572.

Charlesworth D, Charlesworth B, Marais G. 2005. Steps in the evolution of heteromorphic sex chromosomes. *Heredity* **95**: 118-128.

Charlesworth B, Coyne JA, Barton NH. 1987. The relative rates of evolution of sex chromosomes and autosomes. *Am Nat* **130**: 113-146.

Cocquet J, Ellis PJ, Mahadevaiah SK, Affara NA, Vaiman D, Burgoyne PS. 2012. A genetic basis for a postmeiotic X versus Y chromosome intragenomic conflict in the mouse. *PLoS Genet* **8**: e1002900.

Coil D, Jospin G, Darling AE. 2015. A5-miseq: an updated pipeline to assemble microbial genomes from Illumina MiSeq data. *Bioinformatics* **31**: 587-589.

Connallon T, Singh ND, Clark AG. 2012. Impact of genetic architecture on the relative rates of X versus autosomal adaptive substitution. *Mol Biol Evol* **29**: 1933-1942.

Cortez D, Marin R, Toledo-Flores D, Froidevaux L, Liechti A, Waters PD, Grützner F, Kaessmann H. 2014. Origins and functional evolution of Y chromosomes across mammals. *Nature* **508**: 488-493.

Coyne JA. 1992. Genetics and speciation. *Nature* **355**: 511-515.

Coyne JA. 2018. "Two Rules of Speciation" revisted. *Mol Ecol* **27**: 3749-3752.

Coyne JA, Orr HA. 1989. Two rules of speciation. In Otte D, Endler J (eds.) Speciation and its consequences (pp. 180–207). Sinauer, Sunderland, MA.

Coyne JA, Orr HA. 2004. Speciation. Sinauer, Sunderland, MA.

Darwin, C. 1859. On the Origin of Species by Means of Natural Selection or the Preservation of Favored Races in the Struggle for Life. J. Murray. Hachette, UK.

Davis BW, Raudsepp T, Pearks Wilkerson AJ, Agarwala R, Schäffer AA, Houck M, Chowdhary BP, Murphy WJ. 2009. A high-resolution cat radiation hybrid and integrated FISH mapping resource for phylogenomic studies across Felidae. *Genomics* **93**: 299-304.

Davis BW, Seabury CM, Brashear WA, Li G, Roelke-Parker M, Murphy WJ. 2015. Mechanisms Underlying Mammalian Hybrid Sterility in Two Feline Interspecies Models. *Mol Biol Evol* **32**: 2534-2546.

Deakin JE, Koina E, Waters PD, Doherty R, Patel VS, Delbridge ML, Dobson B, Fong J, Hu Y, van den Hurk C et al. 2008. Physical map of two tammar wallaby chromosomes: A strategy for mapping in non-model mammals. *Chromosome Research* **16**: 1159-1175.

Dean MD, Ardlie KG, Nachman MW. 2006. The frequency of multiple paternity suggests that sperm competition is common in house mice (*Mus domesticus*). *Mol Ecol* **15**: 4141-4151.

Delcher AL, Phillippy A, Carlton J, Salzberg SL. 2002. Fast algorithms for large-scale genome alignment and comparison. *Nucl Acid Res* **30**: 2478-2483.

Delgado CLR, Waters PD, Gilbert C, Robinson TJ, Graves JAM. 2009. Physical mapping of the elephant X chromosome: conservation of gene order over 105 million years. *Chromosome Res* **17**: 917-926.

DeLuca DS, Levin JZ, Sivachenko A, Fennell T, Nazaire MD, Williams C, Reich M, Winckler W, Getz G. 2012. RNA-SeQC: RNA-seq metrics for quality control and process optimization. *Bioinformatics* **28**: 1530-1532.

Deng X, Berletch JB, Ngyuen DK, Disteche CM. 2014. X chromosome regulation: diverse patterns in development, tissues and disease. *Nat Rev Genet* **15**: 367-378.

Deng X, Ma W, Ramani V, Hill A, Yang F, Ay F, Berletch JB, Blau CA, Shendure J, Duan Z, Noble WS, Disteche CM. 2015. Bipartite structure of the inactive mouse X chromosome. *Genome Biol* **16**: 152.

Dixon JR, Gorkin DU, Ren B. 2016. Chromatin domains: the units of chromosome organization. *Mol Cell* **62**:668-680.

Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, Hu M, Liu JS, Ren B. 2012. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**: 376.

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**: 15-21.

Dobzhansky T. 1936. Studies on hybrid sterility II. Localization of sterility factors in *Drosophila pseudoobscura* hybrids. *Genetics* **21**: 113-135.

Dod B, Jermilin LS, Boursot P, Chapman VH, Nielsen JT, Bonhomme F. 1993. Counterselection on sex chromosomes in the *Mus musculus* European hybrid zone. *Jour Evol Biol* **6**: 529-546.

Dutheil JY, Munch K, Nam K, Mailund T, Schierup MH. 2015. Strong Selective Sweeps on the X Chromosome in the Human-Chimpanzee Ancestor Explain Its Low Divergence. *PLoS Genet* **11**: e1005451.

Eichler EE, Clark RA, She X. 2004. An assessment of the sequence gaps: Unfinished business in a finished human genome. *Nat Rev Genet* **5**: 345-354.

Eizirik E, Murphy WJ, Koepfli K-P, Johnson WE, Dragoo JW, Wayne RK, O'Brien SJ. 2010. Pattern and timing of diversification of the mammalian order Carnivora inferred from multiple nuclear gene sequences. *Mol Phylogenet Evol* **56**: 49-63.

Ekblom R, Galindo J. 2010. Applications of next generation sequencing in molecular ecology of non-model organisms. *Heredity* **107**: 1-15.

Ellegren H. 2011. Sex-chromosome evolution: recent progress and the influence of male and female heterogamety. *Nat Rev Genet* **12**: 157-166.

Ellegren, H. Genome sequencing and population genomics in non-model organisms. *Trends Ecol* **29**: 51-63.

Emerson JJ, Kaessmann H, Bertrán E, Long M. 2004. Extensive gene traffic on the mammalian X chromosome. *Science* **303**: 537-540.

Engreitz JM, Pandya-Jones A, McDonel P, Shishkin A, Sirokman K, Surka C, Kadri S, Xing J, Goren A, Lander ES et al. 2013. The Xist lncRNA Exploits Three-Dimensional Genome Architecture to Spread Across the X Chromosome. *Science* **341**.

Finestra TR, Gribnau J. 2017. X chromosome inactivation: silencing, topology and reactivation. *Curr Opin Cell Biol* **46**: 54-61.

Fernández AI, Muñoz M, Alves E, Folch JM, Noguera JL, Enciso MP, Rodríguez MDC, Silió L. 2014. Recombination of the porcine X chromosome: a high density linkage map. *BMC genetics* **15**: 148-148.

Galupa R, Heard E. 2017. Topologically Associating Domains in Chromosome Architecture and Gene Regulatory Landscapes during Development, Disease, and Evolution. *Cold Spring Harbor Symposia on Quantitative Biology* **82**: 267-278.

Garrigan D, Kingan SB, Geneva AJ, Vedanayagam JP, Presgraves DC. 2014. Genome diversity and divergence in *Drosophila mauritiana*: multiple signatures of faster X evolution. *Genome Biol Evol* **6**: 2444-2458.

Gel B, Serra E. 2017. karyoploteR: an R/Bioconductor package to plot customizable genomes displaying arbitrary data. *Bioinformatics* **33**: 3088-3090.

Gershony LC, Penedo MC, Davis BW, Murphy WJ, Helps CR, Lyons LA. 2014. Who's behind that mask and cape? The Asian leopard cat's Agouti (ASIP) allele likely affects coat colour phenotype in the Bengal cat breed. *Anim Genet* **45**: 893-897.

Ghenu AH, Bolker BM, Melnick DJ, Evans BJ. 2016. Multicopy gene family evolution on primate Y chromosomes. *BMC Genomics* **17**: 157.

Good JM. 2012. The conflict within and the escalating war between the sex chromosomes. *PLoS Genet* **8**: e1002955.

Good JM, Giger T, Dean MD, Nachman MW. 2010. Widespread overexpression of the X chromosome in sterile F1 hybrid mice. *PLoS Genet* **6**:e101148.

Gordon D, Green P. 2013. Consed: a graphical editor for next-generation sequencing. *Bioinformatics* **29**: 2936-2937.

Graves JAM. 2006. Evolution of vertebrate sex chromosomes and dosage compensation. *Nat Rev Genet* **17**: 33-46.

Graves JA. 2016. Did sex chromosome turnover promote divergence of the major mammal groups? De novo sex chromosomes and drastic rearrangements may have posed reproductive barriers between monotremes, marsupials and placental mammals. *BioEssays* **38**: 734-743.

Graves JAM. 1995. The evolution of mammalian sex chromosomes and the origin of sex determining genes. *Phil Trans R Soc B* **350**: 305-312.

Gray AP. 1972. *Mammalian Hybrids: A Check-list with Bibliography*. Commonwealth Agricultural Bureaux, Farnham Royal, Buckinghamshire.

Haldane JBS. 1922. Sex ratio and unisexual sterility in hybrid animals. *J Genet* **12**: 101-109.

Hallast P, Maisano Delser P, Batini C, Zadik D, Rocchi M, Schempp W, Tyler-Smith C, Jobling MA. 2016. Great-ape Y Chromosome and mitochondrial DNA phylogenies reflect subspecies structure and patterns of mating and dispersal. *Genome Res* **26**:427-439.

Handley LL, Ceplitis H, Ellegren H. Evolutionary strata on the chicken Z chromosome: implications for sex chromosome evolution. *Genetics* **167**: 367-376.

Harris RS. 2007. ‘‘Improved pairwise alignment of genomic DNA.’’ PhD thesis, The Pennsylvania State University.

Hastings PJ, Lupski JR, Rosenberg SM, Ira G. 2009. Mechanisms of change in gene copy number. *Nat Rev Genet* **10**: 551.

Helleu Q, Gérard PR, Montchamp-Moreau C. 2014. Sex Chromosome Drive. *Cold Spring Harb Perspect Biol* **7**: a017616.

Hobza R, Kejnovsky E, Vyskot B, Widmer A. 2007. The role of chromosomal rearrangements in the evolution of *Silene latifolia* sex chromosomes. *Mol Genet Genomics* **278**: 633-638.

Hu Y, Wu Q, Ma S, Ma T, Shan L, Wang X, Nie Y, Ning Z, Yan L, Xiu Y et al. 2017. Comparative genomics reveals convergent evolution between the bamboo-eating giant and red pandas. *Proc Natl Acad Sci U.S.A.* **114**: 1081-1086.

Huddleston J, Ranade S, Malig M, Antonacci F, Chaisson M, Hon L, Sudmant PH, Graves TA, Alkan C, Dennis MY et al. 2014. Reconstructing complex regions of genomes using long-read sequencing technology. *Genome Res* **24**: 688-696.

Hughes JF, Page DC. 2015. The biology and evolution of mammalian Y chromosomes. *Annu Rev Genet* **49**: 507-527.

Hughes JF, Skaletsky H, Brown LG, Pyntikova T, Graves T, Fulton RS, Dugan S, Ding Y, Buhay CJ, Kremitzki C et al. 2012. Strict evolutionary conservation followed rapid gene loss on human and rhesus Y chromosomes. *Nature* **483**: 82-86.

Hughes JF, Skaletsky H, Pyntikova T, Graves TA, van Daalen SK, Minx PJ, Fulton RS, McGrath SD, Locke DP, Friedman C et al. 2010. Chimpanzee and human Y

chromosomes are remarkably divergent in structure and gene content. *Nature* **463**: 536-539.

Irwin, DE. 2018. Sex chromosomes and speciation in birds and other ZW systems. *Mol Ecol* **27**: 3831-3851.

Janečka JE, Davis BW, Ghosh S, Paria N, Das PJ, Orlando L, Schubert M, Nielsen MK, Stout TAE, Brashear W et al. 2018. Horse Y chromosome assembly displays unique evolutionary features and putative stallion fertility genes. *Nat Commun* **9**: 2945.

Jégu T, Aeby E, Lee JT. 2017. The X chromosome in space. *Nat Rev Genet* **18**: 377-339.

Jobling MA, Tyler-Smith C. 2017. Human Y-chromosome variation in the genome-sequencing era. *Nat Rev Genet* **18**: 485-497.

Johnson NA. 2000. Speciation: Dobzhansky-Muller incompatibilities, dominance and gene interactions. *Trends Ecol* **15**: 480-482.

Johnson WE, Eizirik E, Pecon-Slattery J, Murphy WJ, Antunes A, Teeling E, O'Brien SJ. 2006. The late Miocene radiation of modern Felidae: a genetic assessment. *Science* **311**: 73-77.

Johnson WE, Onorato DP, Roelke ME, Land ED, Cunningham M, Belden RC, McBride R, Jansen D, Lotz M, Shindle D et al. 2010. Genetic restoration of the Florida panther. *Science* **329**: 1641-1645.

Karmin M, Saag L, Vicente M, Wilson Sayres MA, Järve M, Talas UG, Rootsi S, Ilumäe AM, Mägi R, Mitt M et al. 2015. A recent bottleneck of Y chromosome diversity coincides with a global change in culture. *Genome Res* **25**: 459-466.

Katoh K, Standley DM. 2013. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Mol Biol Evol* **30**: 772-780.

Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C et al. 2012. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28**: 1647-1649.

Kent WJ. 2002. BLAT—The BLAST-Like Alignment Tool. *Genome Res* **12**: 656-664.

Khil PP, Smirnova NA, Romanienko PJ, Camerini-Otero RD. 2004. The mouse X chromosome is enriched for sex-biased genes not subject to selection by meiotic sex chromosome inactivation. *Nat Genet* **36**: 642-646.

Khost DE, Eickbush DG, Larracuente AM. 2017. Single-molecule sequencing resolves the detailed structure of complex satellite DNA loci in *Drosophila melanogaster*. *Genome Res* **27**:709-721.

Kirkpatrick M, Hall DW. 2004. Male-biased mutation, sex linkage, and the rate of adaptive evolution. *Evolution* **58**: 437–440.

Kitchener AC, Breitenmoser-Würsten C, Eizirik E, Gentry A, Werdelin L, Wilting A, Yamaguchi N, Abramov AV, Christiansen P, Driscoll C et al. 2017. A revised taxonomy of the Felidae: The final report of the Cat Classification Task Force of the IUCN/SSC Cat Specialist Group. *Cat News Special Issue* **11**: 1-80.

Kong A, Gudbjartsson DF, Sainz J, Jonsdottir GM, Gudjonsson SA, Richardsson B, Sigurdardottir S, Barnard J, Hallbeck B, Masson G. 2002. A high-resolution recombination map of the human genome. *Nat genet* **31**: 241.

Koren S, Harhay GP, Smith TP, Bono JL, Harhay DM, Mcvey SD, Radune D, Bergman NH, Phillippy AM. 2013. Reducing assembly complexity of microbial genomes with single-molecule sequencing. *Genome Biol* **14**: R101.

Koren S, Schatz MC, Walenz BP, Martin J, Howard JT, Ganapathy G, Wang Z, Rasko DA, McCombie WR, Jarvis ED et al. 2012. Hybrid error correction and de novo assembly of single-molecule sequencing reads. *Nat Biotechnol* **30**: 693-700.

Koren S, Walenz BP, Berlin K, Miller JR, Phillippy AM. 2017. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Res* **27**: 722-736.

Kuroda-Kawaguchi T, Skaletsky H, Brown LG, Minx PJ, Cordum HS, Waterston RH, Wilson RK, Silber S, Oates R, Rozen S et al. 2001. The AZFc region of the Y chromosome features massive palindromes and uniform recurrent deletions in infertile men. *Nat Genet* **29**: 279-286.

Krefting J, Andrade-Navarro MA, Ibn-Salem J. 2018. Evolutionary stability of topologically associating domains is associated with conserved gene regulation. *BMC Biol* **16**: 87.

Kuroiwa A, Tsuchiya K, Watanabe T, Hishigaki H, Takahashi E, Namikawa T, Matuda Y. 2001. Conservation of the rat X chromosome gene order in rodent species. *Chromosome Res* **9**: 61-67.

Lahn BT, Page DC. 1997. Functional coherence of the human Y chromosome. *Science* **278**: 675-680.

Lahn BT, Page DC. 1999. Four evolutionary strata on the human X chromosome. *Science* **286**: 964-967.

Lamichhaney S, Berglund J, Almén MS, Maqbool K, Grabherr M, Martinez-Barrio A, Promerová M, Rubin CJ, Wang C, Zamani N et al. 2015. Evolution of Darwin's finches and their beaks revealed by genome sequencing. *Nature* **518**: 371-375.

Lange J, Skaletsky H, van Daalen SKM, Embry SL, Korver CM, Brown LG, Oates RD, Silber S, Repping S, Page DC. 2009 Isodicentric Y chromosomes and sex disorders as byproducts of homologous recombination that maintains palindromes. *Cell* **138**: 855-869.

Langmead B, Trapnell C, Pop M, Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10**: R25.

Larson EL, Keeble S, Vanderpool D, Dean MD, Good JM. 2016. The composite regulatory basis of the large X-effect in mouse speciation. *Mol Biol Evol* **34**: 282-295.

Laurie, CC. 1997. The weaker sex is heterogametic: 75 years of Haldane's Rule. *Genetics* **147**: 937-951.

Leggett RM, Clavijo BJ, Clissold L, Clark MD, Caccamo M. 2014. NextClip: an analysis and read preparation tool for Nextera Long Mate Pair libraries. *Bioinformatics* **30**: 566-568.

Lercher MJ, Urratia AO, Hurst LD. 2003. Evidence that the human X chromosome is enriched for male-specific but not female-specific genes. *Mol Biol Evol* **20**:1113-1116.

Li G, Davis BW, Eizirik E, Murphy WJ. 2016. Phylogenomic evidence for ancient hybridization in the genomes of living cats (Felidae). *Genome Res* **26**: 1-11.

Li G, Davis BW, Raudsepp T, Pearks Wilkerson AJ, Mason VC, Ferguson-Smith M, O'Brien PC, Waters PD, Murphy WJ. 2013. Comparative analysis of mammalian Y chromosomes illuminates ancestral structure and lineage-specific evolution. *Genome Res* **23**: 1486-1495.

Li G, Figueiro HV, Eizirik E, Murphy WJ. 2018. Recombination-aware phylogenomics unravels the complex divergence of hybridizing species. *bioRxiv* doi:10.1101/485904: 485904.

Li G, Hillier LW, Grahn RA, Zimin AV, David VA, Menotti-Raymond M, Middleton R, Hannah S, Hendrickson S, Makunin A et al. 2016. A high-resolution SNP array-based linkage map anchors a new domestic cat draft genome assembly and provides detailed patterns of recombination. *G3* **6**:1607-1616.

Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**: 1754-1760.

Li H, Wysoker A, Handsaker B, Marth G, Abecasis G, Ruan J, Homer N, Durbin R, Fennell T. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**: 2078-2079.

Lindblad-Toh K, Garber M, Zuk O, Lin MF, Parker BJ, Washietl S, Kheradpour P, Ernst J, Jordan G, Mauceli E et al. 2011. A high-resolution map of human evolutionary constraint using 29 mammals. *Nature* **478**: 476-482.

Lonsdale J, Thomas J, Salvatore M, Phillips R, Lo E, Shad S, Hasz R, Walters G, Garcia F, Young N et al. 2013. The Genotype-Tissue Expression (GTEx) project. *Nature Genet* **45**: 580-585.

Lucotte EA, Skov L, Jensen JM, Macià MC, Munch K, Schierup MH. 2018. Dynamic Copy Number Evolution of X- and Y-Linked Ampliconic Genes in Human Populations. *Genetics* **209**: 907.

Lupiáñez Darío G, Kraft K, Heinrich V, Krawitz P, Brancati F, Klopocki E, Horn D, Kayserili H, Opitz John M, Laxova R et al. 2015. Disruptions of Topological Chromatin Domains Cause Pathogenic Rewiring of Gene-Enhancer Interactions. *Cell* **161**: 1012-1025.

Lyon MF. 1961. Gene action in the X-chromosome of the mouse (*Mus musculus* L.). *Nature* **190**: 372-373.

Lyon, MF. 1992. Some milestones in the history of X-chromosome inactivation. *Annu Rev Genet* **26**: 17-29.

Ma J, Iannuccelli N, Duan Y, Huang W, Guo B, Riquet J, Huang L, Milan D. 2010. Recombinational landscape of porcine X chromosome and individual variation in female meiotic recombination associated with haplotypes of Chinese pigs. *BMC Genomics* **11**: 159.

MacDonald DQ, Loveridge AJ. 2010. Biology and conservation of wild felids. Oxford University Press, Oxford.

Mank JE, Vicoso B, Berlin S, Charlesworth B. 2009. Effective population size and the faster-X effect: empirical results and their interpretation. *Evolution* **64**: 663-674.

Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J* **17**: dx.doi.org/10.14806/ej.17.1.200.

McLaughlin RN, Malik HS. 2017. Genetic conflicts: the usual suspects and beyond. *J Exp Biol* **220**: 6-17.

Meadows JRS, Lindblad-Toh K. 2017. Dissecting evolution and disease using comparative vertebrate genomics. *Nat Rev Genet* **18**: 624-636.

Meiklejohn CD, Tao Y. 2010. Genetic conflict and sex chromosome evolution. *Trends Ecol Evol* **25**: 215-223.

Meisel RP, Connallon T. 2013. The faster-X effect: integrating theory and data. **29**: 537-544.

Melé M, Ferreira PG, Reverter F, DeLuca DS, Monlong J, Sammeth M, Young TR, Goldmann JM, Pervouchine DD, Sullivan TJ et al. 2015. Human genomics. The human transcriptome across tissues and individuals. *Science* **348**: 660-665.

Mihola O, Trachtulec Z, Vlcek C, Schimenti JC, Forejt J. 2009. A Mouse Speciation Gene Encodes a Meiotic Histone H3 Methyltransferase. *Science* **323**: 373-375.

Montague MJ, Li G, Gandolfi B, Khan R, Aken BL, Searle SM, Minx P, Hillier LW, Koboldt DC, Davis BW et al. 2014. Comparative analysis of the domestic cat genome reveals genetic signatures underlying feline biology and domestication. *Proc Natl Acad Sci* **111**: 17230-17235.

Morgan TH, Bridges CB. 1916. Sex-linked inheritance in *Drosophila*. Carnegie Institution of Washington, Washington.

Mueller JL, Mahadevaiah SK, Park PJ, Warburton PE, Page DC, Turner JMA. 2008. The mouse X chromosome is enriched for multi-copy testis genes exhibiting post-meiotic expression. *Nat Genet* **40**:794-799.

Mueller JL, Skaletsky H, Brown LG, Zaghlul S, Rock S, Graves T, Auger K, Warren WC, Wilson RK, Page DC. 2013. Independent specialization of the human and mouse X chromosomes for the male germ line. *Nat Genet* **45**: 1083-1087.

Muller HJ. 1942. Isolating mechanisms, evolution, and temperature. *Biol Symp* **6**: 71-125.

Murphy WJ, Eizirik E, Johnson WE, Zhang YP, Ryder OA, O'Brien SJ. 2001. Molecular phylogenetics and the origin of placental mammals. *Nature* **614**: 614-618.

Murphy WJ, Larkin DM, der Wind AE-v, Bourque G, Tesler G, Auvil L, Beever JE, Chowdhary BP, Galibert F, Gatzke L et al. 2005. Dynamics of Mammalian

Chromosome Evolution Inferred from Multispecies Comparative Maps. *Science* **309**: 613.

Murphy WJ, Sun S, Chen Z, Pecon-Slattery J, O'Brien SJ. 1999. Extensive conservation of sex chromosome organization between cat and human revealed by parallel radiation hybrid mapping. *Genome Res* **9**: 1223-1230.

Murphy WJ, Wilkerson AJP, Raudsepp T, Agarwala R, Schäffer AA, Stanyon R, Chowdhary BP. 2006. Novel gene acquisition on carnivore Y chromosomes. *PLoS Genet* **2**: e43.

Myers EW, Sutton GG, Delcher AL, Dew IM, Fasulo DP, Flanigan MJ, Kravitz SA, Mobarry CM, Reinert KHJ, Remington KA et al. 2000. A Whole-Genome Assembly of *Drosophila*. *Science* **287**: 2196.

Nadeau JH. 1989. Maps of linkage and synteny homologies between mouse and man. *Trends in Genetics* **5**: 82-86.

Nagaraja R, MacMillan S, Kere J, Jones C, Griffin S, Schmatz M, Terrell J, Shomaker M, Jermak C, Hott C, et al. 1997. X chromosome map at 75-kb STS resolution, revealing extremes of recombination and GC content. *Genome Res* **7**: 210-222.

Nam K, Munch K, Hobolth A, Dutheil JY, Veeramah KR, Woerner AE, Hammer MF, Great Ape Genome Diversity P, Mailund T, Schierup MH. 2015. Extreme selective sweeps independently targeted the X chromosomes of the great apes. *Proc Nat Assoc Sci* **112**: 6413-6418.

Necsulea A, Soumillon M, Warnefors M, Liechti A, Daish T, Zeller U, Baker JC, Grützner F, Kaessmann H. 2014. The evolution of lncRNA repertoires and expression patterns in tetrapods. *Nature* **505**: 635-640.

Nesterova TB, Duthie SM, Mazurok NA, Isaenko AA, Rubstova NV, Zakian SM, Brockdorff N. 1998. Comparative mapping of X chromosomes in vole species of the genus *Microtus*. *Chromosome Res* **6**:41-48.

Noor MAF, Feder JL. 2006. Speciation genetics: evolving approaches. *Nat Rev Genet* **7**: 851-861.

Nora EP, Lajoie BR, Schulz EG, Giorgetti L, Okamoto I, Servant N, Piolot T, van Berkum NL, Meisig J, Sedat J et al. 2012. Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* **485**: 381.

Nosil P, Feder JL. 2012. Genomic divergence during speciation: causes and consequences. *Phil Trans R Soc B* **367**: 332-342.

Oetjens MT, Shen F, Emery SB, Zou Z, Kidd JM. 2016. Y-Chromosome Structural Diversity in the Bonobo and Chimpanzee Lineages. *Genome Biol Evol* **8**: 2231-2240.

Ohno S. 1967. Sex chromosomes and sex-linked genes. Springer-Verlag.

Ohno S, Hauschka TS. 1960. Allocycly of the X-chromosome in tumors and normal tissues. *Cancer Res* **20**: 541-545.

Orr HA. 1997. Haldane's rule. *Annu Rev Ecol Evol Syst* **28**: 195–218.

Palmer S, Perry J, Ashworth A. 1995. A contravention of Ohno's law in mice. *Nat Genet* **10**: 472–476.

Paria N, Raudsepp T, Pearks Wilkerson AJ, O'Brien PC, Ferguson-Smith MA, Love CC, Arnold C, Rakestraw P, Murphy WJ, Chowdhary BP. 2011. A gene catalogue of the euchromatic male-specific region of the horse Y chromosome: comparison with human and other mammals. *PLOS ONE* **6**: e21374.

Park W, Oh HS, Kim H. 2002. Acceleration of X-chromosome gene order evolution in the cattle lineage. *BMB Rep* **46**: 310-315.

Paudel Y, Madsen O, Megens HJ, Frantz LA, Bosse M, Crooijmans RP, Groenen MA. 2015. Copy number variation in the speciation of pigs: a possible prominent role for olfactory receptors. *BMC Genomics* **16**: 330. doi: 10.1186/s12864-015-1449-9.

Payseur BA, Hoekstra HE. 2006. Signatures of reproductive isolation in patterns of single nucleotide diversity across inbred strains of mice. *Genetics* **171**: 1905-1916.

Pearks Wilkerson AJ, Raudsepp T, Graves T, Albracht D, Warren W, Chowdhary BP, Skow LC, Murphy WJ. 2008. Gene discovery and comparative analysis of X-degenerate genes from the domestic cat Y chromosome. *Genomics* **92**: 329-338.

Penny GD, Kay GF, Sheardown SA, Rastan S, Brockdorff N. 1996. Requirement for *Xist* in X chromosome inactivation. *Nature* **379**: 131-137.

Pevzner P, Tesler G. 2003a. Genome Rearrangements in Mammalian Evolution: Lessons From Human and Mouse Genomes. *Genome Res* **13**: 37-45.

Pevzner P, Tesler G. 2003b. Human and mouse genomic sequences reveal extensive breakpoint reuse in mammalian evolution. *Proc Natl Acad Sci* **100**: 7672-7677.

Piumi F, Schibler L, Vaiman D, Oustry A, Cribiu EP. 1998. Comparative cytogenetic mapping reveals chromosome rearrangements between the X chromosomes of two

closely related mammalian species (cattle and goats). *Cytogenet Cell Genet* **81**: 36-41.

Plath K, Mlynarczyk-Evans S, Nusinow DA, Panning B. 2002. Xist RNA and the Mechanism of X Chromosome Inactivation. Annu Rev Genet **36**: 233-278.

Potrzebowski L, Vinckenbosch N, Marques AC, Chalmel F, Jégou B, Kaessman H. 2008. Chromosomal gene movements reflect the recent origin and biology of therian sex chromosomes. *PLoS Biol* **6**: e80. doi:10.1371/journal.pbio.0060080.

Poznik GD, Xue Y, Mendez FL, Willems TF, Massaia A, Wilson Sayres MA, Ayub Q, McCarthy SA, Narechania A, Kashin S et al. 2016. Punctuated bursts in human male demography inferred from 1,244 worldwide Y-chromosome sequences. *Nat Genet* **48**: 593-599.

Prakash B, Kuosku V, Olsaker I, Gustavsson I, Chowdhary BP. 1996. Comparative FISH mapping of bovine cosmids to reindeer chromosomes demonstrates conservation of the X-chromosome. *Chromosome Res* **4**: 214-217.

Presgraves DC. 2008. Sex chromosomes and speciation in *Drosophila*. *Trends Genet* **24**: 336-343.

Presgraves DC. 2018. Evaluating genomic signatures of "the large X-effect" during complex speciation. *Mol Ecol* **27**: 3822-3830.

Pukazhenthi BS, Neubauer K, Jewgenow K, Howard J, Wildt DE. 2006. The impact and potential etiology of teratospermia in the domestic cat and its wild relatives. *Theriogenology* **66**: 112-121.

Pukazhenthi BS, Wildt DE, Howard JG. 2001. The phenomenon and significance of teratospermia in felids. *J Reprod Fertil Suppl* **57**: 423-433.

Quilter CR, Blott SC, Mileham AJ, Affara NA, Sargent CA, Griffin DK. 2002. A mapping and evolutionary study of porcine sex chromosome genes. *Mamm Genome* **13**: 588-594.

Qvarnström A, Bailey RI. 2008. Speciation through evolution of sex-linked genes. *Heredity* **102**: 4-15.

Ramirez F, Bhardwaj V, Arrigoni L, Lam KC, Gruning BA, Villaveces J, Habermann B, Akhtar A, Manke T. 2018. High-resolution TADs reveal DNA sequences underlying genome organization in flies. *Nat Comm* doi: 10.1038/s41467-017-02525-w.

Raudsepp T, Lee EJ, Kata SR, Brinkmeyer C, Mickelson JR, Skow LC, Womack JE, Chowdhary B. 2004. Exceptional conservation of horse-human gene order on X chromosome revealed by high-resolution radiation hybrid mapping. *Proc Nat Assoc Sci* **101**: 2386-2391.

Raudsepp T, Chowdhary BP. 2008. FISH for mapping single copy genes. *Methods Mol Biol* **422**: 31-49.

Repping S, Skaletsky H, Lange J, Silber S, Van Der Veen F, Oates RD, Page DC, Rozen S. 2002. Recombination between palindromes P5 and P1 on the human Y chromosome causes massive deletions and spermatogenic failure. *Am J Hum Genet* **71**: 906-922.

Rice WR. 1984. Sex chromosomes and the evolution of sexual dimorphism. *Evolution* **4**: 735-742.

Rice WR. 1987. Genetic hitchhiking and the evolution of reduced genetic activity of the Y sex chromosome. *Genetics* **116**: 161-167.

Rice WR. 1996. Evolution of the Y sex chromosome in animals: Y chromosomes evolve through the degeneration of autosomes. *BioSci* **5**: 331-343.

Richards S, Liu Y, Bettencourt BR, Hradecky P, Letovsky S, Nielsen R, Thornton K, Hubisz MJ, Chen R, Meisel RP et al. 2005. Comparative genome sequencing of Drosophila pseudoobscura: Chromosomal, gene, and cis-element evolution. *Genome Res* **15**: 1-18.

Robinson TJ, Harrison WR, Ponce de León FA, Davis SK, Elder FFB. 1998. A molecular cytogenetic analysis of X chromosome repatterning in the Bovidae: transpositions, inversions, and phylogenetic inference. *Cytogenet Cell Genet* **80**: 179-184.

Roelke ME, Martenson JS, O'Brien SJ. 1993. The consequences of demographic reduction and genetic depletion in the endangered Florida panther. *Curr Biol* **3**: 340-350.

Ross MT, Grafham DV, Coffey AJ, Scherer S, McLay K, Muzny D, Platzer M, Howell GR, Burrows C, Bird CP et al. 2005. The DNA sequence of the human X chromosome. *Nature* **434**: 325-337.

Rozen S, Skaletsky H, Marszalek JD, Minx PJ, Cordum HS, Waterston RH, Wilson RK, Page DC. 2003. Abundant gene conversion between arms of palindromes in human and ape Y chromosomes. *Nature* **423**: 873-876.

Rugarli EI, Adler DA, Borsani G, Tsuchiya K, Franco B, Hauge X, Disteche C, Chapman V, Ballabio A. 1995. Different chromosomal localization of the Clcn4 gene in *Mus spretus* and C57BL/6J mice. *Nat Genet* **10**:466–471.

Sandstedt SA, Tucker PK. 2004. Evolutionary strata on the mouse chromosome correspond to strata on the human X chromosome. *Genome Res* **14**: 267-272.

Schilthuizen M, Giesbers MCWG, Beukeboom LW. 2011. Haldane's rule in the 21st century. *Heredity* **107**: 95-102.

Schmittgen TD, Livak KJ. 2008. Analyzing real-time PCR data by the comparative CT method. *Nat Protoc* **3**: 1101-1108.

Schubert S, Dechend F, Skawran B, Krawczak M, Schmidtke J. Molecular evolution of the murine *tspy* genes. *Cytogenet Cell Genet* **91**: 239-242.

Seehausen O, Butlin RK, Keller I, Wagner CE, Boughman JW, Hohenlohe PA, Peischel CL, Saetre GP, Bank C, Brannstorm A. 2014. Genomics and the origin of species. *Nat Rev Genet* **15**: 176-192.

Skaletsky H, Kuroda-Kawaguchi T, Minx PJ, Cordum HS, Hillier L, Brown LG, Repping S, Pyntikova T, Ali J, Bieri T et al. 2003. The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. *Nature* **423**: 825-837.

Simecek P, Forejt J, Williams RW, Shiroishi T, Takada T, Lu L, Johnson TE, Bennett B, Deschepper CF, Scott-Boyer M-P et al. 2017. High-Resolution Maps of Mouse Reference Populations. *G3* **7**: 3427.

Skinner BM, Sargent CA, Churcher C, Hunt T, Herrero J, Loveland JE, Dunn M, Louzada S, Fu B, Chow W et al. 2016. The pig X and Y Chromosomes: structure, sequence, and evolution. *Genome Res* **26**: 130-139.

Smit, AFA, Hubley, R & Green, P. RepeatMasker Open-4.0. 2013-2015 http://www.repeatmasker.org

Soh YQ, Alföldi J, Pyntikova T, Brown LG, Graves T, Minx PJ, Fulton RS, Kremitzki C, Koutseva N, Mueller JL et al. 2014. Sequencing the mouse Y chromosome reveals convergent gene acquisition and amplification on both sex chromosomes. *Cell* **159**: 800-813.

Soumillon M, Necsulea A, Weier M, Brawand D, Zhang X, Gu H, Barthès P, Kokkinaki M, Nef S, Gnirke A et al. 2013. Cellular source and mechanisms of high transcriptome complexity in the mammalian testis. *Cell Rep* **3**: 2179-2190.

Spencer JA, Sinclair AH, Watson JM, Graves JAM. 1991a. Genes on the short arm of the human X chromosome are not shared with the marsupial X. *Genomics* **11**: 339-345.

Spencer JA, Watson JM, Graves JAM. 1991b. The X chromosome of marsupials shares a highly conserved region with eutherians. *Genomics* **9**: 598-604.

Spriggs HF, Holmes NG, Breen M, Deloukas P, Langford CF, Ross MT, Carter NP, Davis ME, Knights C, Smith A, et al. 2003. Construction and integration of radiation-hybrid and cytogenetic maps of dog Chromosome X. *Mamm Genome* **14**: 214-221.

Stamatakis A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* **22**: 2688-2690.

Sudmant PH, Rausch T, Gardner EJ, Handsaker RE, Abyzov A, Huddleston J, Zhang Y, Ye K, Jun G, Hsi-Yang Fritz M et al. 2015. An integrated map of structural variation in 2,504 human genomes. *Nature* **526**: 75.

Tao Y, Masly JP, Araripe L, Ke Y, Hartl DL. 2007. A sex-ratio meiotic drive system in Drosophila simulans. I: an autosomal suppressor. *PLoS Biology* **5**: e292.

Terrell KA, Crosier AE, Wildt DE, O'Brien SJ, Anthony NM, Marker L, Johnson WE. 2016. Continued decline in genetic diversity among wild cheetahs (Acinonyx jubatus) without further loss of semen quality. *Biol Cons* **200**: 192-199.

Toder R, O'Neill RJW, Wienberg J, O'Brien P, Voullaire L, Graves JAM. 1997. Comparative chromosome painting between two marsupials: origings of an XX/XY1Y2 sex chromosome system. *Mamm Genome* **8**: 418-422.

Tomaszkiewicz M, Rangavittal S, Cechova M, Campos Sanchez R, Fescemyer HW, Harris R, Ye D, O'Brien PC, Chikhi R, Ryder OA et al. 2016. A time- and cost-effective strategy to sequence mammalian Y Chromosomes: an application to the de novo assembly of gorilla Y. *Genome Res* **26**: 530-540.

Touré A, Clemente EJ, Ellis PJ, Mahadevaiah SK, Ojarikre OA, Ball PA, Reynard L, Loveland KL, Burgoyne PS, Affara NA. 2005. Identification of novel Y chromosome encoded transcripts by testis transcriptome analysis of mice with deletions of the Y chromosome long arm. *Genome Biol* **6**: R102.

Touré A, Grigoriev V, Mahadevaiah SK, Rattigan A, Ojarikre OA, Burgoyne PS. 2004. A protein encoded by a member of the multicopy Ssty gene family located on the long arm of the mouse Y chromosome is expressed during sperm development. *Genomics* **83**: 140-147.

Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, Pimentel H, Salzberg SL, Rinn JL, Pachter L. 2012. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat Proto* **7**: 562-578.

Treangen TJ, Sommer DD, Angly FE, Koren S, Pop M. 2011. Next Generation Sequence Assembly with AMOS. *Curr Protoc Bioinformatics* doi:10.1002/0471250953.bi1108s33.

Trigo TC, Freitas TRO, Kunzler G, Cardoso L, Silva JCR, Johnson WE, O'Brien SJ, Bonatto SL, Eizirik E. 2008. Inter-species hybridization among Neotropical cats of the genus Leopardus, and evidence for an introgressive hybrid zone between L. geoffroyi and L. tigrinus in southern Brazil. *Molec Ecol* **17**: 4317-4333.

Trigo Tatiane C, Schneider A, de Oliveira Tadeu G, Lehugeur Livia M, Silveira L, Freitas Thales RO, Eizirik E. 2013. Molecular Data Reveal Complex Hybridization and a Cryptic Species of Neotropical Wild Cat. *Curr Biol* **23**: 2528-2533.

Trigo TC, Tirelli FP, de Freitas TRO, Eizirik E. 2014. Comparative Assessment of Genetic and Morphological Variation at an Extensive Hybrid Zone between Two Wild Cats in Southern Brazil. *PLoS ONE* **9**: e108469.

Turner JMA. 2007. Meiotic sex chromosome inactivation. *Development* **134**: 1823-1831.

Turelli M, Orr HA. 2000. Dominance, epistasis, and the genetics of postzygotic isolation. *Genetics* **154**: 1663-1679.

Veyrunes F, Waters PD, Miethke P, Rens W, McMillian D, Alsop AE, Grützner F, Deakin JE, Whittington CM, Schatzkamer K, et al. 2008. Bird-like sex chromosomes of platypus imply recent origin of mammal sex chromosomes. *Genome Res* **18**: 965-973.

Vicoso B, Charlesworth B. 2006. Evolution on the X chromosome: unusual patterns and processes. *Nat Rev Genet* **7**: 645-653.

Vicoso B, Charlesworth B. 2009. Effective population size and the faster-X effect: an extended model. *Evolution* **63**: 2413–2426.

Vinckenbosch N, Dupanloup I, Kaessmann H. 2006. Evolutionary fate of retroposed gene copies in the human genome. *Proc Natl Acad Sci* **103**: 3220-3225.

Wang P, McCarrey JR, Yang F, Page DC. 2001. An abundance of X-linked genes expressed in spermatogonia. *Nat Genet* **27**: 422-426.

Wang X, Douglas KC, VandeBerg JL, Clark AG, Samollow PB. 2014. Chromosome-wide profiling of X-chromosome inactivation and epigenetic states in fetal brain and placenta of the opossum, Monodelphis domestica. *Genome Res* **24**: 70-83.

Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, Cuomo CA, Zeng Q, Wortman J, Young SK et al. 2014. Pilon: An Integrated Tool for Comprehensive

Microbial Variant Detection and Genome Assembly Improvement. *PLOS ONE* **9**: e112963.

Watanabe TK, Bihoreau MT, McCarthy LC, Kiguwa SL, Hishigaki H, Tsuji A, Browne J, Yamasaki Y, Mizoguchi-Miyakita A, Oga K, et al. 1999. A radiation hybrid map of the rat genome containing 5,255 markers. *Nature* **22**: 27-36.

Waterston RH, Lindblad-Toh K, Birney E, Rogers J, Abril JF, Agarwal P, Agarwala R, Ainscough R, Alexandersson M, An P, et al. Initial sequencing and comparative analysis of the mouse genome. *Nature* **420**: 520-562.

Watson JM, Spencer JA, Graves JAM, Snead ML, Lau EC. 1992. Autosomal localization of the amelogenin gene in monotremes and marsupials: implications for mammalian sex chromosome evolution. *Genomics* **14**: 785-789.

Werren JH. 2011. Selfish genetic elements, genetic conflict, and evolutionary innovation. *Proc Natl Acad Sci* **108**: 10863-10870.

Wildt DE, Phillips LG, Simmons LG, Chakraborty PK, Brown JL, Howard JG, Teare A, Bush M. 1988. A comparative analysis of ejaculate and hormonal characteristics of the captive male cheetah, tiger, leopard, and puma. *Biol Reprod* **38**: 245-255.

Wolf JBW, Lindell J, Backström N. 2010. Speciation genetics: current status and evolving approaches. *Phil Trans R Soc B* **365**: 1717-1733.

Wong AK, Ruhe AL, Dumont BL, Roberson KR, Guerrero G, Shull SM, Ziegle JS, Millon LV, Broman KW, Payseur BA et al. 2010. A comprehensive linkage map of the dog genome. *Genetics* **184**: 595-605.

Wright AE, Dean R, Zimmer F, Mank JE. How to make a sex chromosome. *Nat Comm* **7**: 1-8.

Wu CI, Davis AW. 1993. Evolution of postmating reproductive isolation: the composite nature of Haldane's Rule and its genetic basis. *Am Nat* **142**: 187-212.

Wu CI, Johnson NA, Palopoli MF (1996). Haldane's Rule and its legacy: why are there so many sterile males? *Trends Ecol Evol* **11**: 281–284.

Wu CI. Ting C. Genes and speciation. *Nat Rev Genet* **5**: 114-122.

Wu TD, Watanabe CK. 2005. GMAP: a genomic mapping and alignment program for mRNA and EST sequences. *Bioinformatics* **21**: 1859-1875.

Yang Z. 2007. PAML 4: Phylogenetic Analysis by Maximum Likelihood. *Molec Biol Evol* **24**: 1586-1591.

Zhang G, Li C, Li Q, Li B, Larkin DM, Lee C, Storz JF, Antunes A, Greenwold MJ, Meredith RW et al. 2014. Comparative genomics reveals insights into avian genome evolution and adaptation. *Science* **346**: 1311-1320.

APPENDIX A

SUPPLEMENTAL FIGURES

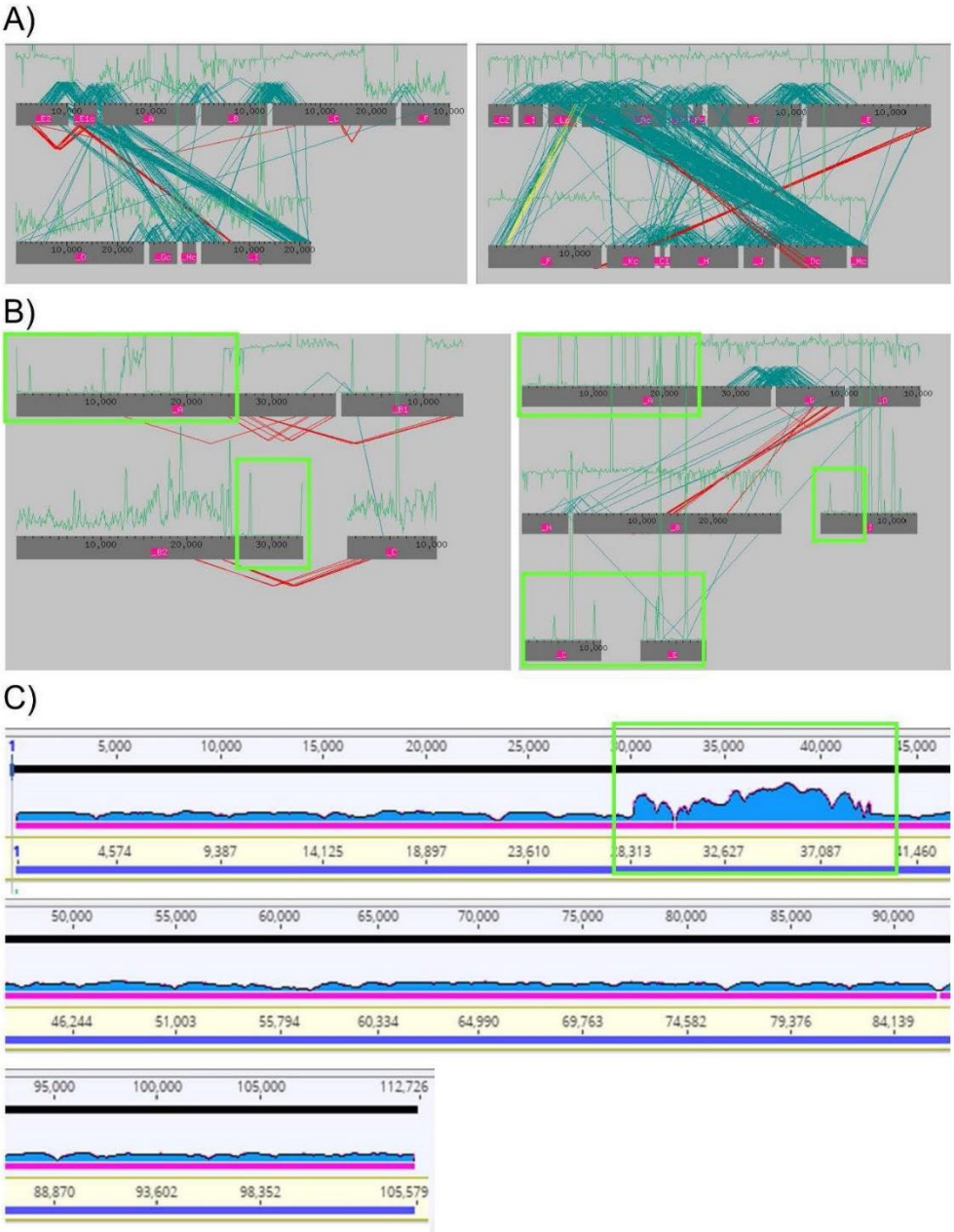


Figure A2.1. Frequently encountered problems when assembling ampliconic regions of the domestic cat Y chromosome using Illumina short read data. (*A*) Initial assemblies were too fragmented to accurately scaffold. Teal lines represent mate pairs from 3 and 8 kb jumping libraries, while the green lines represent read depth of mapped mate pairs. Patterns as shown are likely caused by large segmental duplications with high pairwise identity within the sequenced insert from the selected BAC. (*B*) Large areas of contigs where read depth was drastically reduced (or missing all together) are outlined by green boxes. (*C*) MiSeq libraries were mapped back to assembled scaffolds and examined for disparities in read depth coverage. The sudden and isolated increase in read depth (outlined in green) for a region of this assembly suggests that a large segmental duplication has been collapsed and the actual sequence of the clone has not been correctly reconstructed.

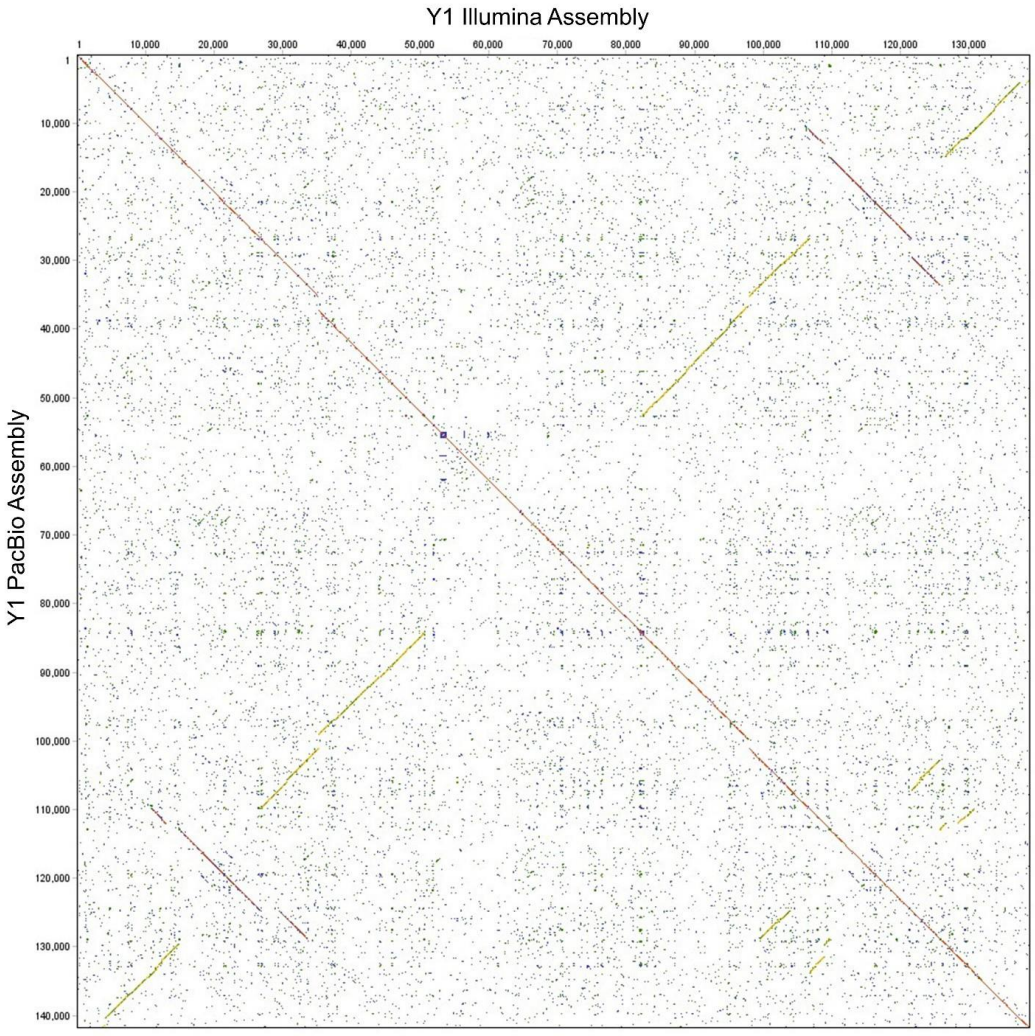


Figure A2.2. Dotplot alignment of two separate assemblies of BAC clone 529O11 using Illumina (top) and PacBio (side) sequencing.

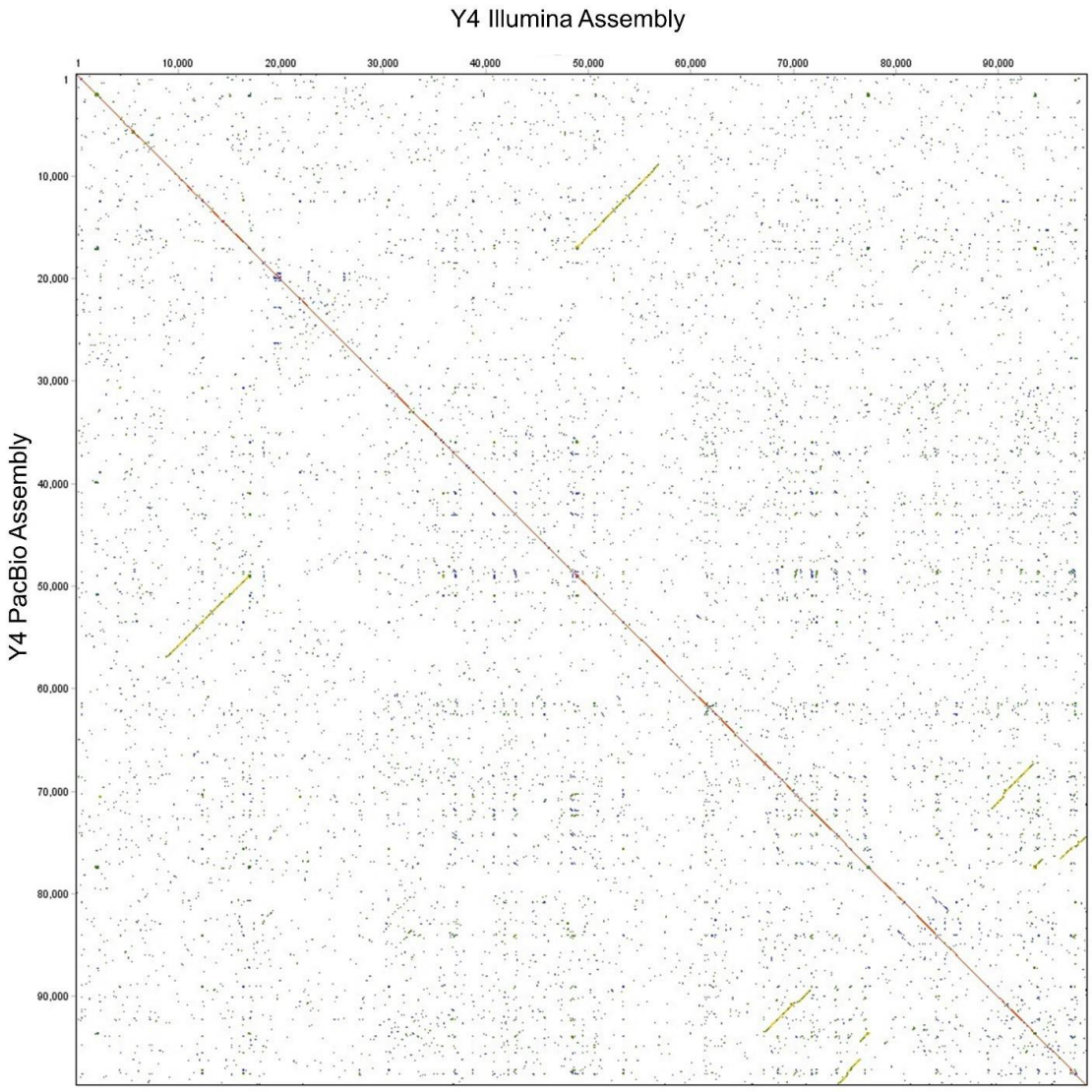


Figure A2.3. Dotplot alignment of two separate assemblies of BAC clone 532F8 using Illumina (top) and PacBio (side) sequencing.

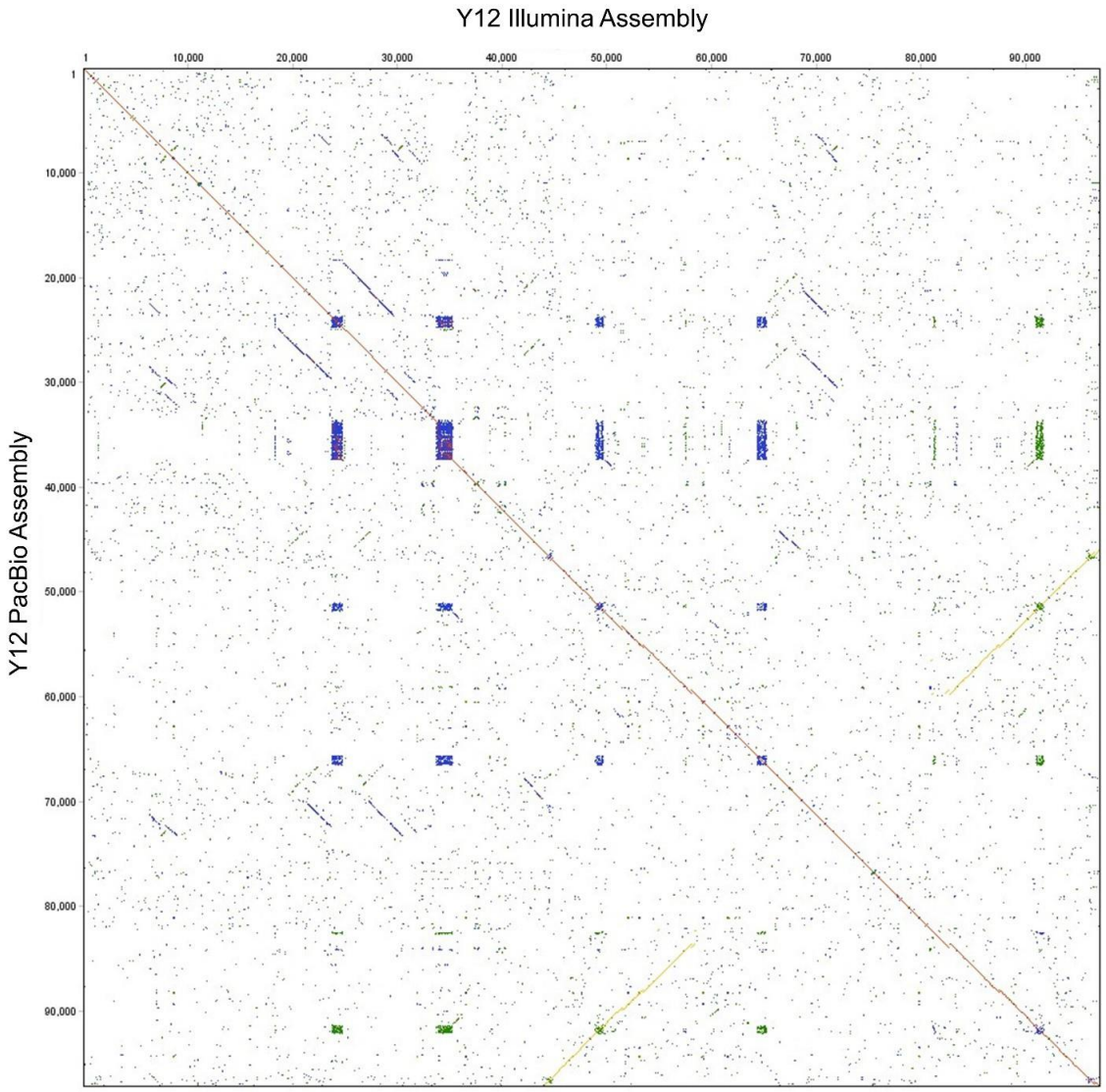


Figure A2.4. Dotplot alignment of two separate assemblies of BAC clone 544B22 using Illumina (top) and PacBio (side) sequencing.

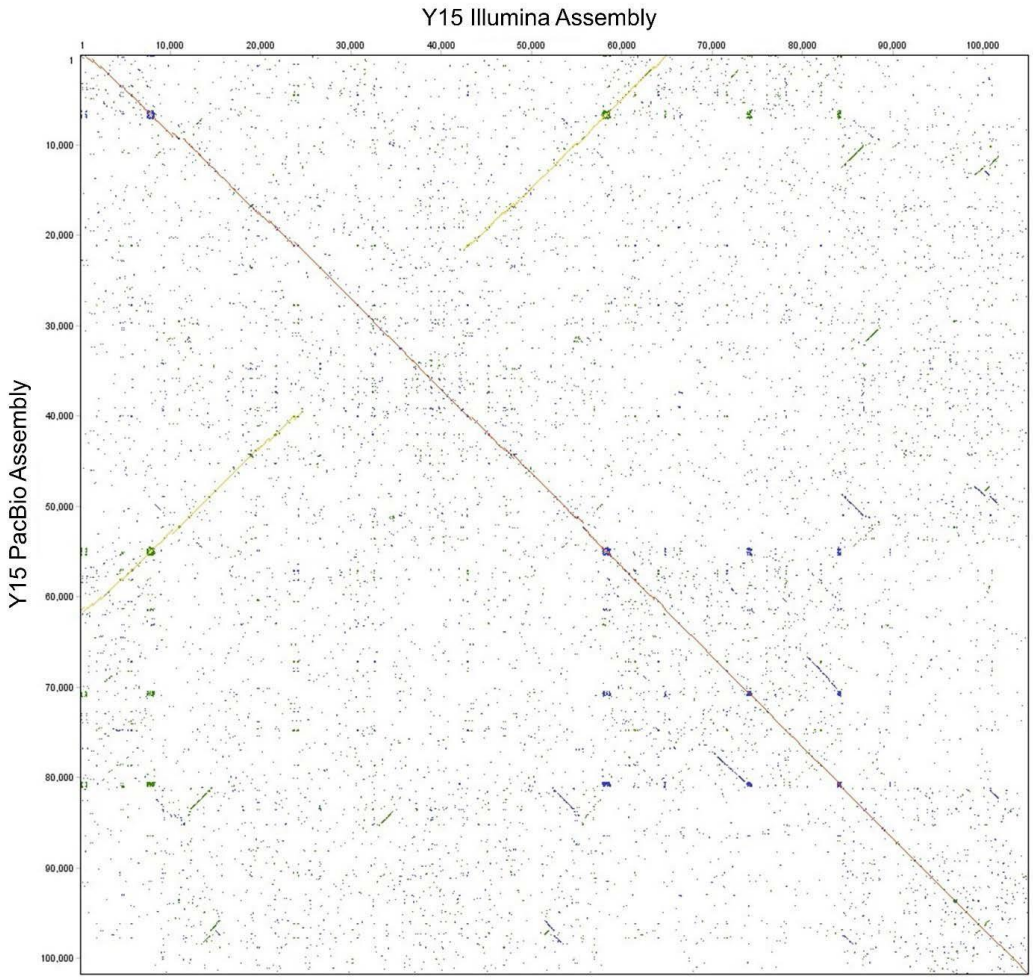


Figure A2.5. Dotplot alignment of two separate assemblies of BAC clone 554L9 using Illumina (top) and PacBio (side) sequencing.

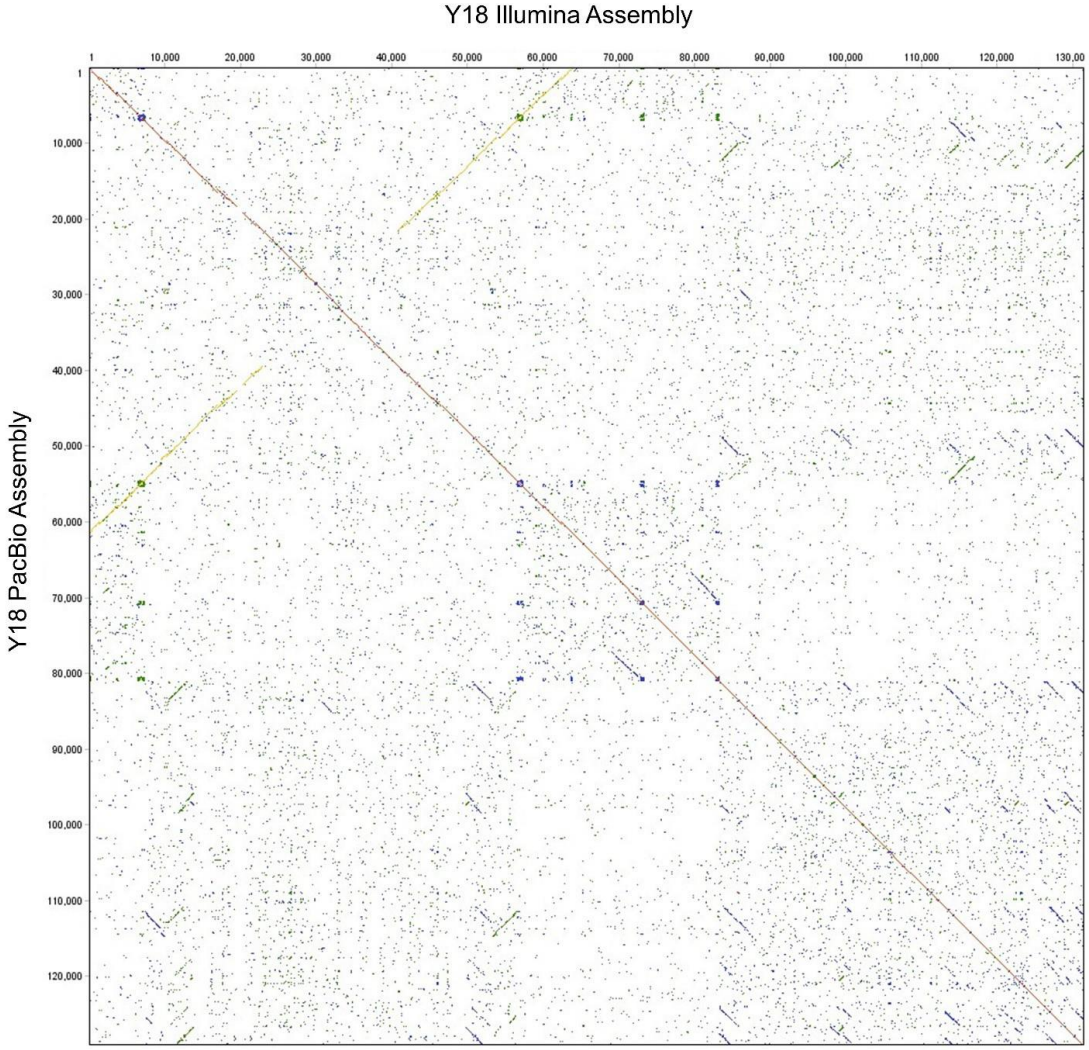


Figure A2.6. Dotplot alignment of two separate assemblies of BAC clone 572E6 using Illumina (top) and PacBio (side) sequencing.

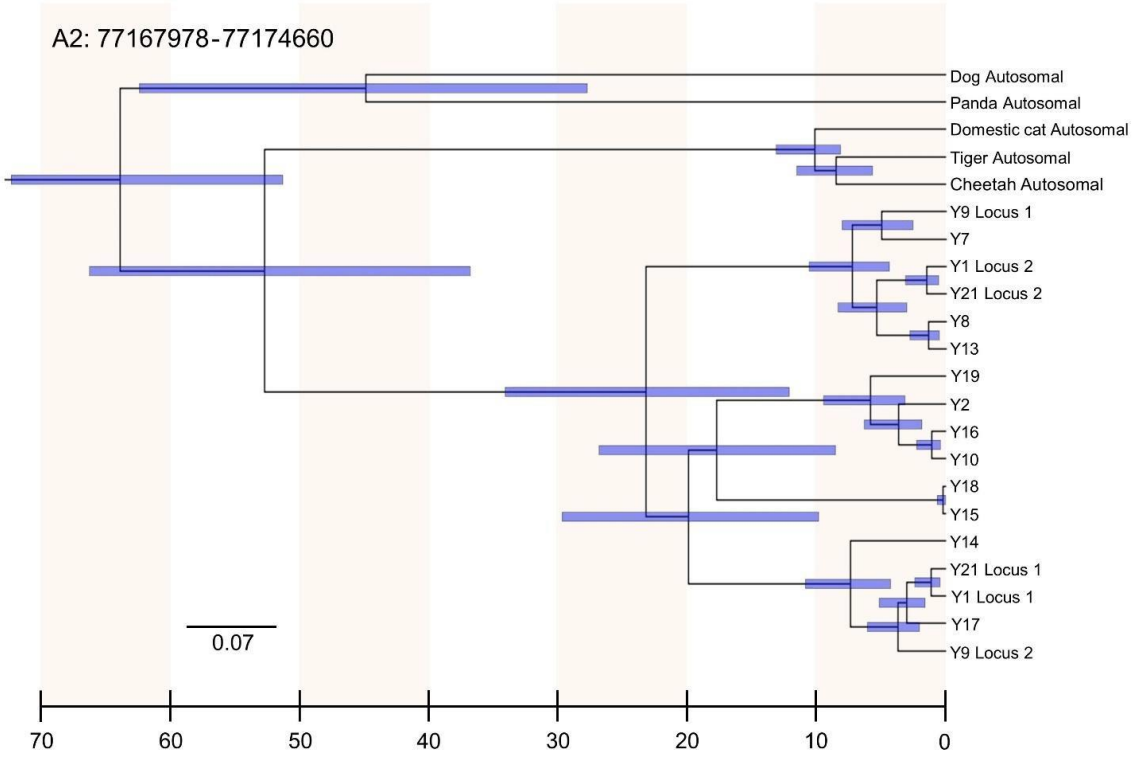


Figure A2.7. Molecular timescale for the transposition of chrA2:77167978- 77174660 to the felid Y chromosome (mya). 95% confidence intervals are depicted as the horizontal blue lines.

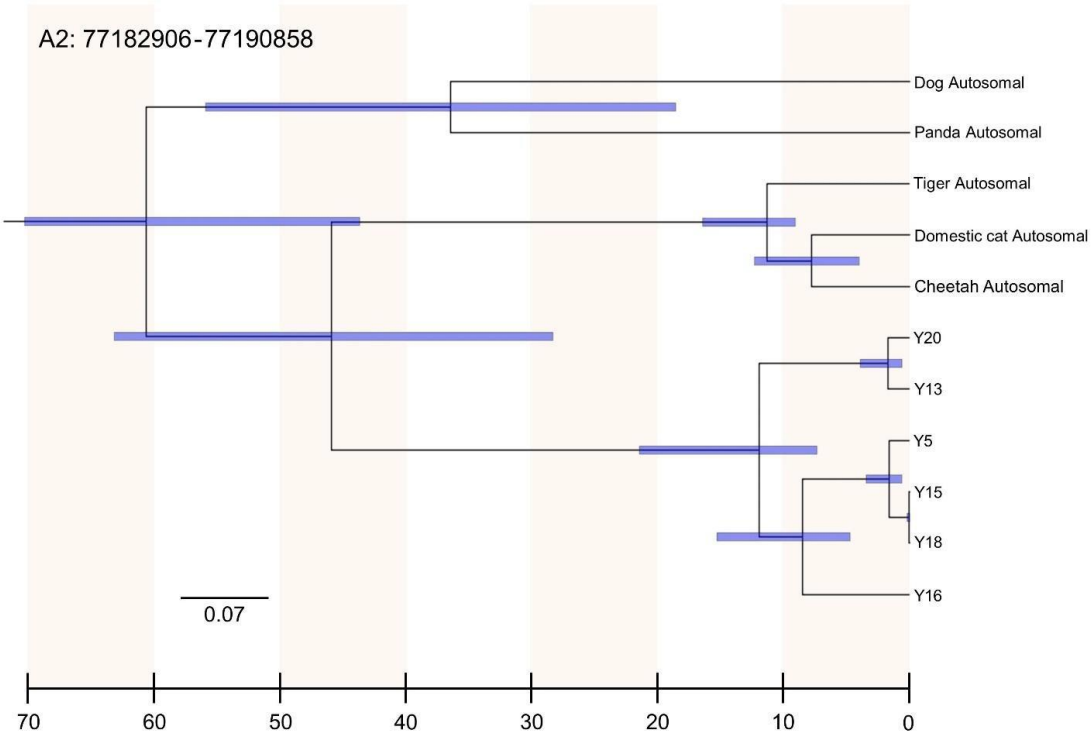


Figure A2.8. Molecular timescale for the transposition of chrA2:77182906- 77190858 to the felid Y chromosome (mya). 95% confidence intervals are depicted as the horizontal blue lines.

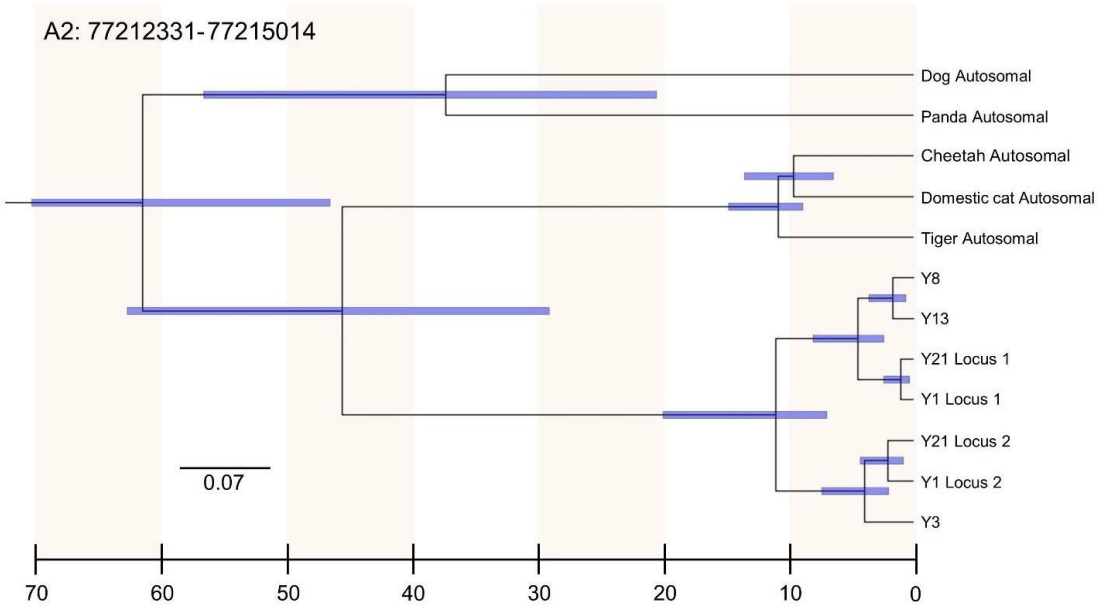


Figure A2.9. Molecular timescale for the transposition of chrA2:77212331- 77215014 to the felid Y chromosome (mya). 95% confidence intervals are depicted as the horizontal blue lines.

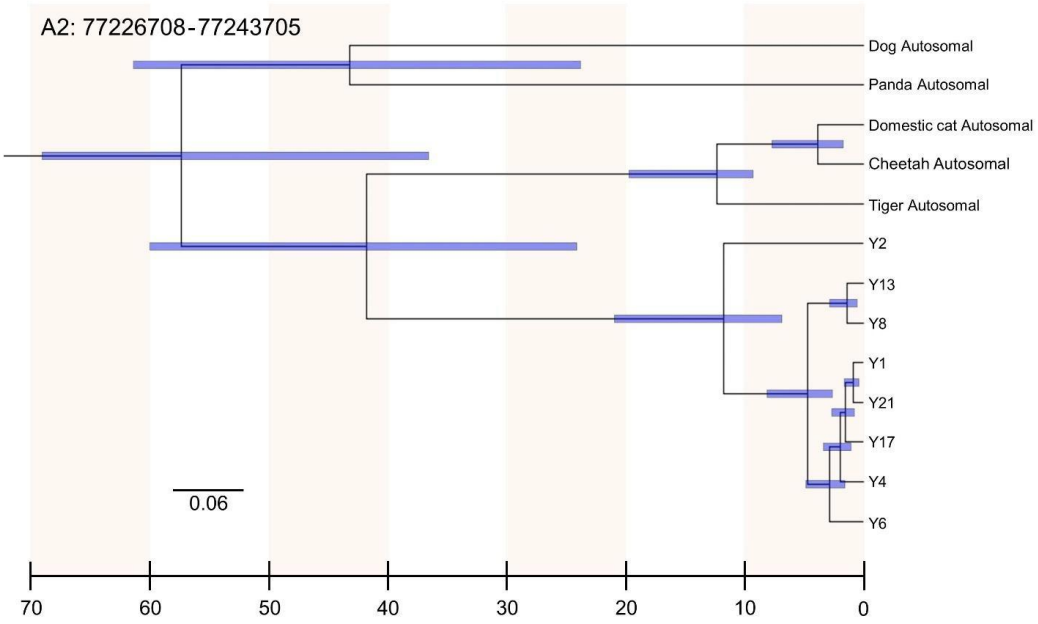


Figure A2.10. Molecular timescale for the transposition of chrA2:77226708- 77243705 to the felid Y chromosome (mya). 95% confidence intervals are depicted as the horizontal blue lines.

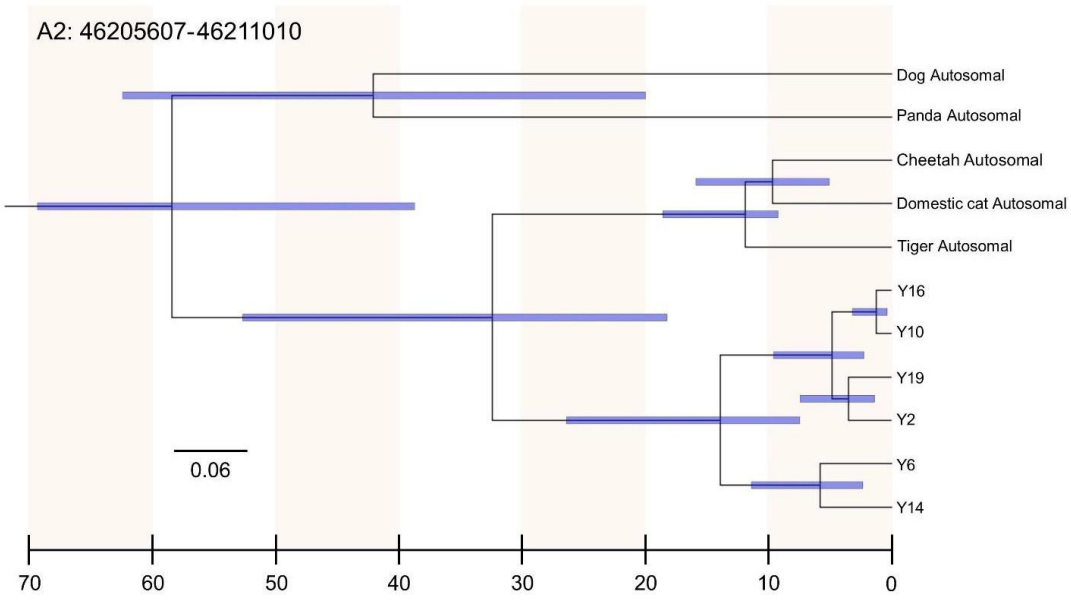


Figure A2.11. Molecular timescale for the transposition of chrA2:46205607- 46211010 to the felid Y chromosome (mya). 95% confidence intervals are depicted as the horizontal blue lines.

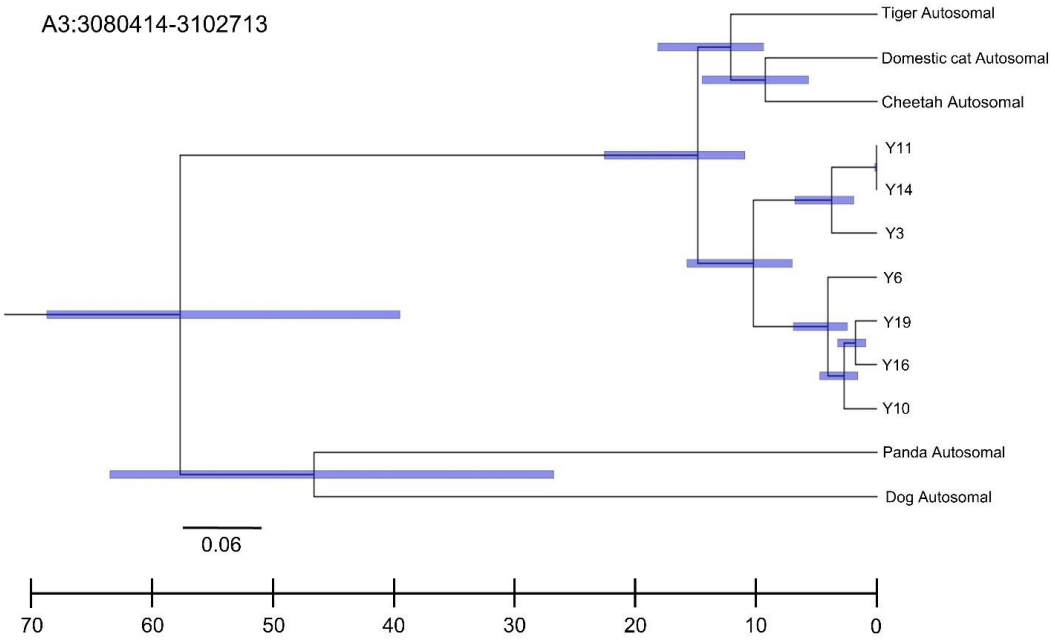


Figure A2.12. Molecular timescale for the transposition of chrA3:3080414- 3102713 to the felid Y chromosome (mya). 95% confidence intervals are depicted as the horizontal blue lines.

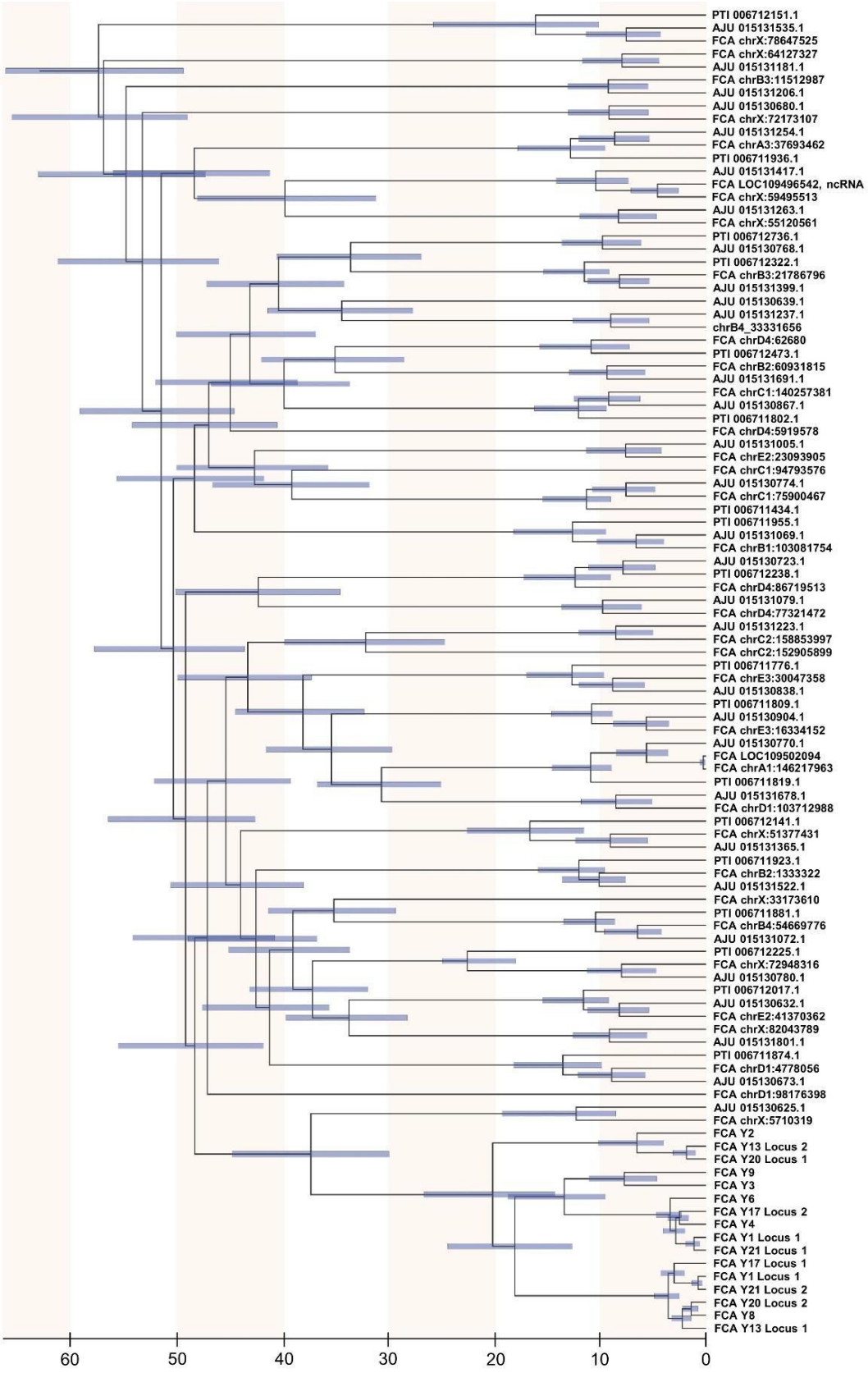

PTI 006712151.1
AJU 015131535.1
FCA chrX:78647525
FCA chrX:64127327
AJU 015131181.1
FCA chrB3:11512987
AJU 015131206.1
AJU 015130680.1
FCA chrX:72173107
AJU 015131254.1
FCA chrA3:37693462
PTI 006711936.1
AJU 015131417.1
FCA LOC109496542, ncRNA
FCA chrX:59495513
AJU 015131263.1
FCA chrX:55120561
PTI 006712736.1
AJU 015130768.1
PTI 006712322.1
FCA chrB3:21786796
AJU 015131399.1
AJU 015130639.1
AJU 015131237.1
chrB4_33331656
FCA chrD4:62680
PTI 006712473.1
FCA chrB2:60931815
AJU 015131691.1
FCA chrC1:140257381
AJU 015130867.1
PTI 006711802.1
FCA chrD4:5919578
AJU 015131005.1
FCA chrE2:23093905
FCA chrC1:94793576
AJU 015130774.1
FCA chrC1:75900467
PTI 006711434.1
PTI 006711955.1
AJU 015131069.1
FCA chrB1:103081754
AJU 015130723.1
PTI 006712238.1
FCA chrD4:86719513
AJU 015131079.1
FCA chrD4:77321472
AJU 015131223.1
FCA chrC2:158853997
FCA chrC2:152905899
PTI 006711776.1
FCA chrE3:30047358
AJU 015130838.1
PTI 006711809.1
AJU 015130904.1
FCA chrE3:16334152
AJU 015130770.1
FCA LOC109502094
FCA chrA1:146217963
PTI 006711819.1
AJU 015131678.1
FCA chrD1:103712988
PTI 006712141.1
FCA chrX:51377431
AJU 015131365.1
PTI 006711923.1
FCA chrB2:1333322
AJU 015131522.1
FCA chrX:33173610
PTI 006711881.1
FCA chrB4:54669776
AJU 015131072.1
PTI 006712225.1
FCA chrX:72948316
AJU 015130780.1
PTI 006712017.1
AJU 015130632.1
FCA chrE2:41370362
FCA chrX:82043789
AJU 015131801.1
PTI 006711874.1
FCA chrD1:4778056
AJU 015130673.1
FCA chrD1:98176398
AJU 015130625.1
FCA chrX:5710319
FCA Y2
FCA Y13 Locus 2
FCA Y20 Locus 1
FCA Y9
FCA Y3
FCA Y6
FCA Y17 Locus 2
FCA Y4
FCA Y1 Locus 1
FCA Y21 Locus 1
FCA Y17 Locus 1
FCA Y1 Locus 1
FCA Y21 Locus 2
FCA Y20 Locus 2
FCA Y8
FCA Y13 Locus 1
60
50
40
30
20
10
0

Figure A2.13. Molecular timescale for the endogenous retrovirus-like element to the felid Y chromosome (mya). 95% confidence intervals are depicted as the horizontal blue lines.

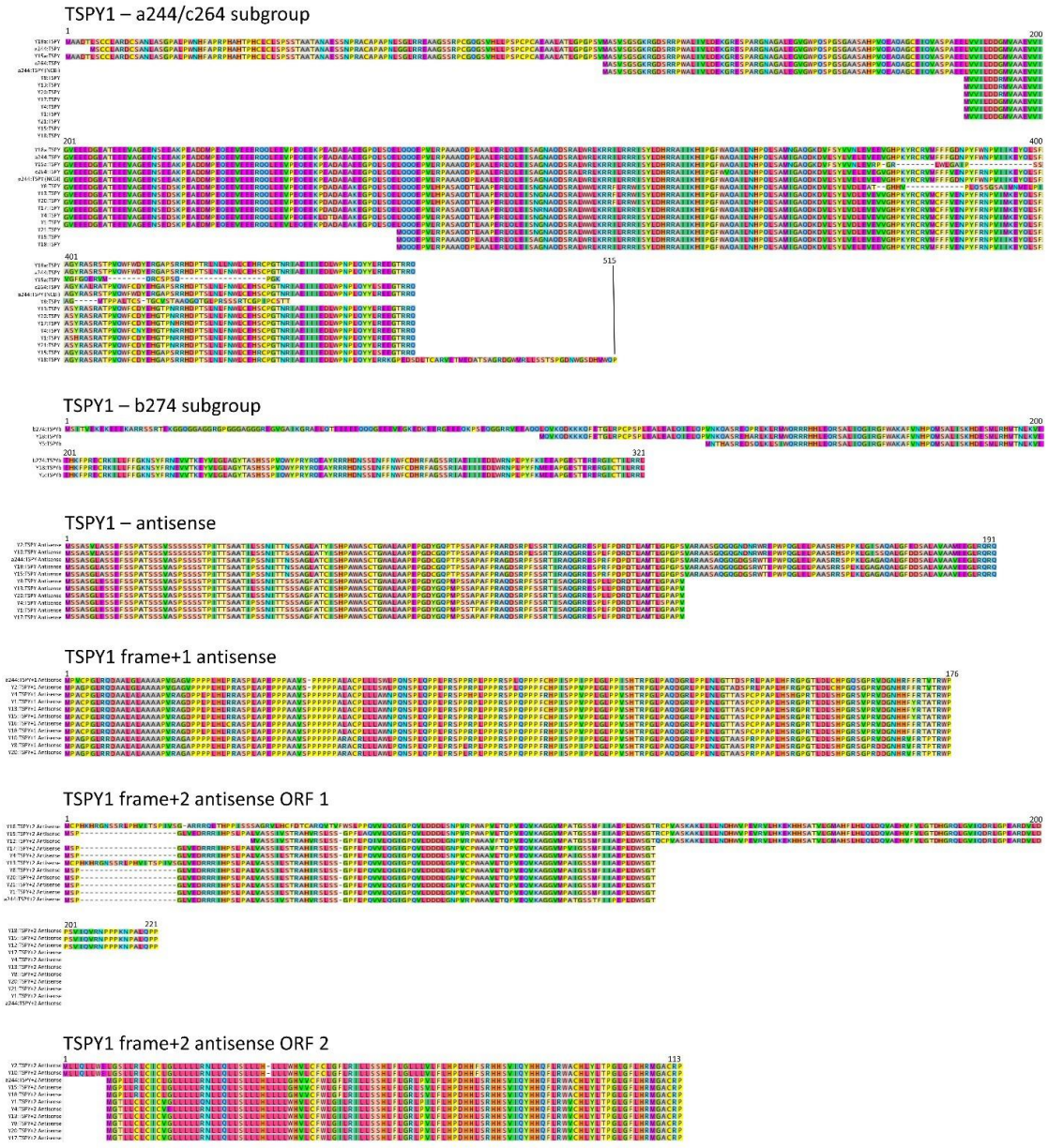


Figure A2.14. Protein alignments of open reading frames found in *TSPY1* variants identified in this study.

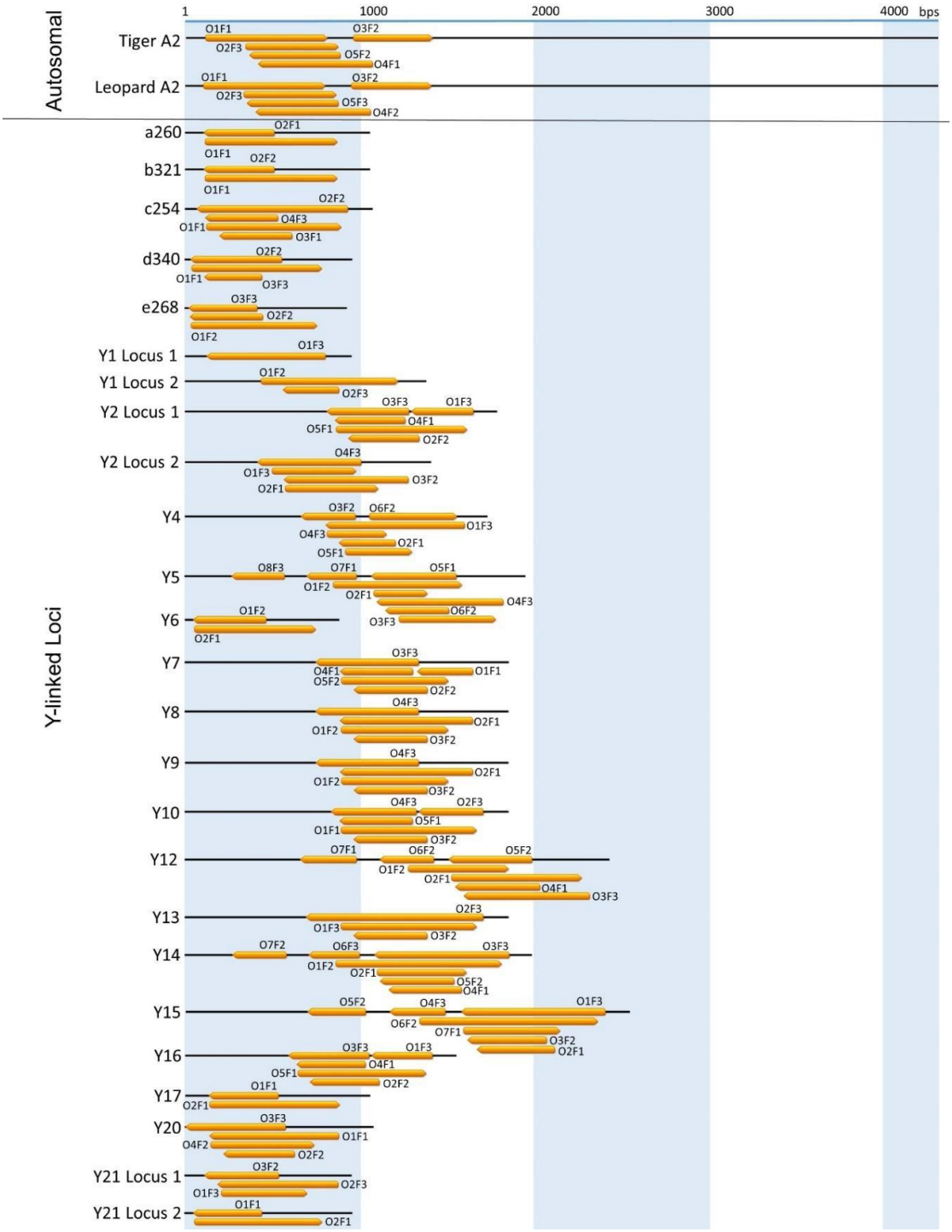


Figure A2.15. Unaligned sequences from CCDC71L and CCDC71LY loci depicting the open reading frames used for protein alignments.

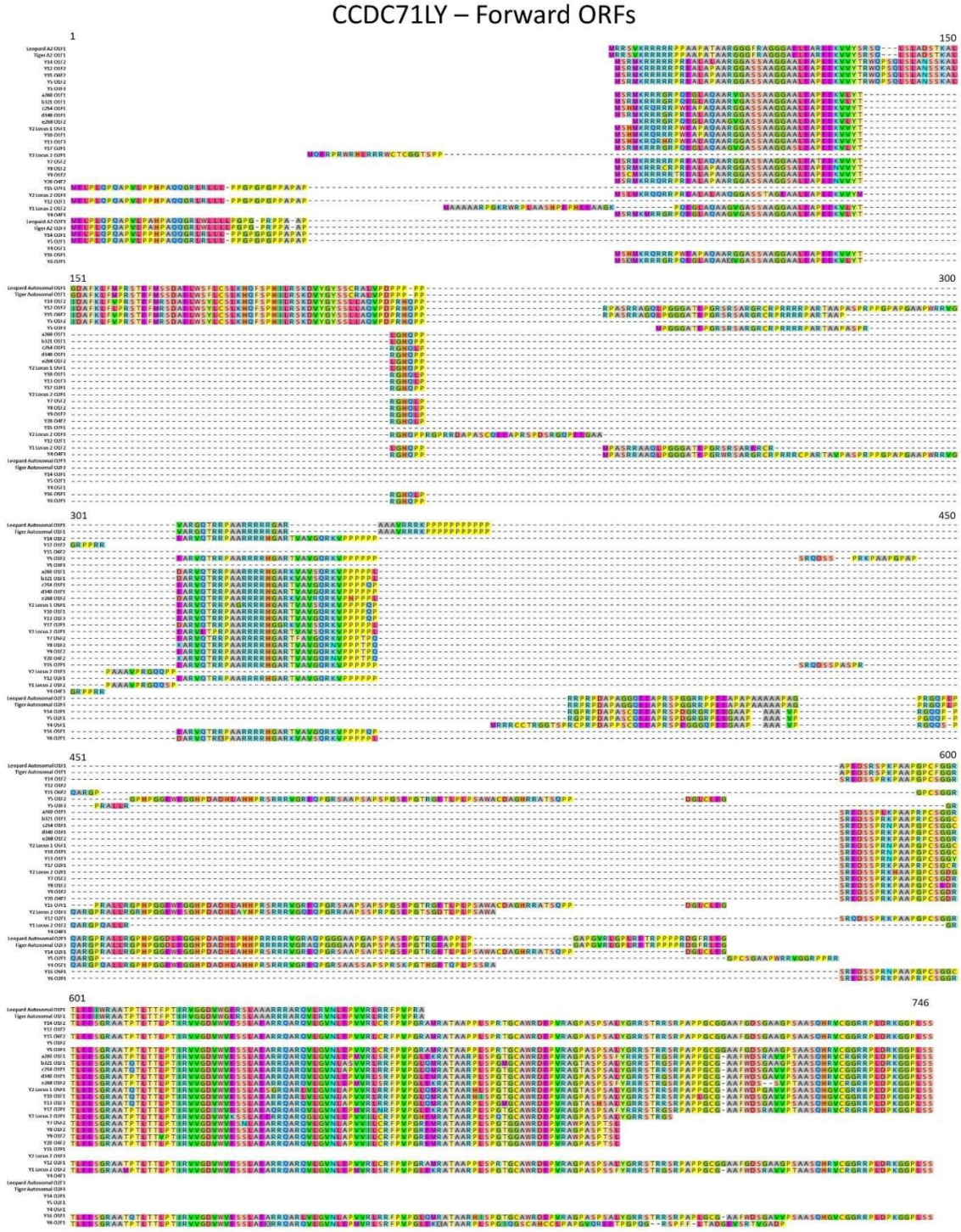


Figure A2.16. Protein alignments from the forward open reading frames of *CCDC71L*, previously published *CCDC71LY* transcripts, and the Y-linked *CCDC71LY* loci identified in this study.

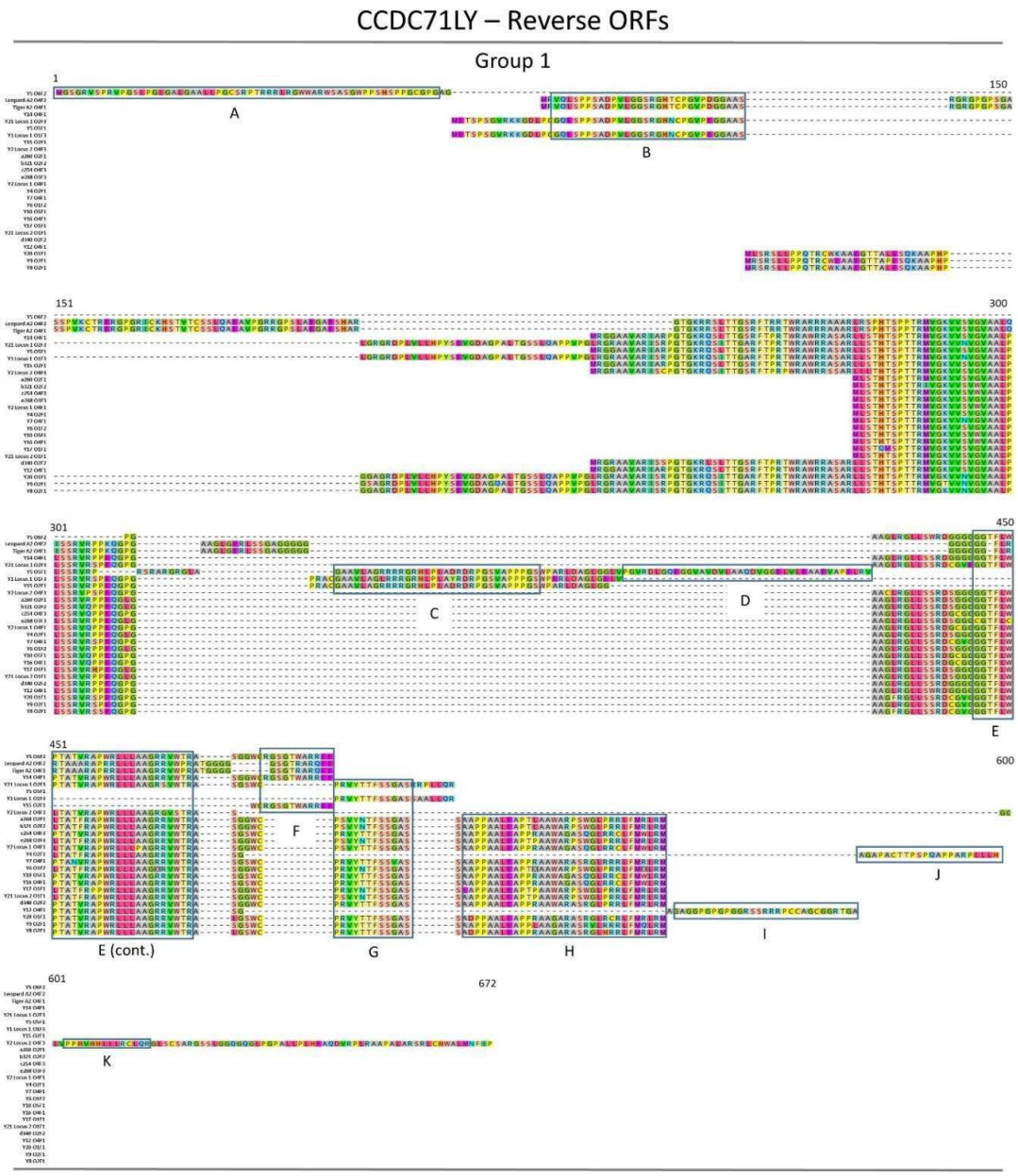
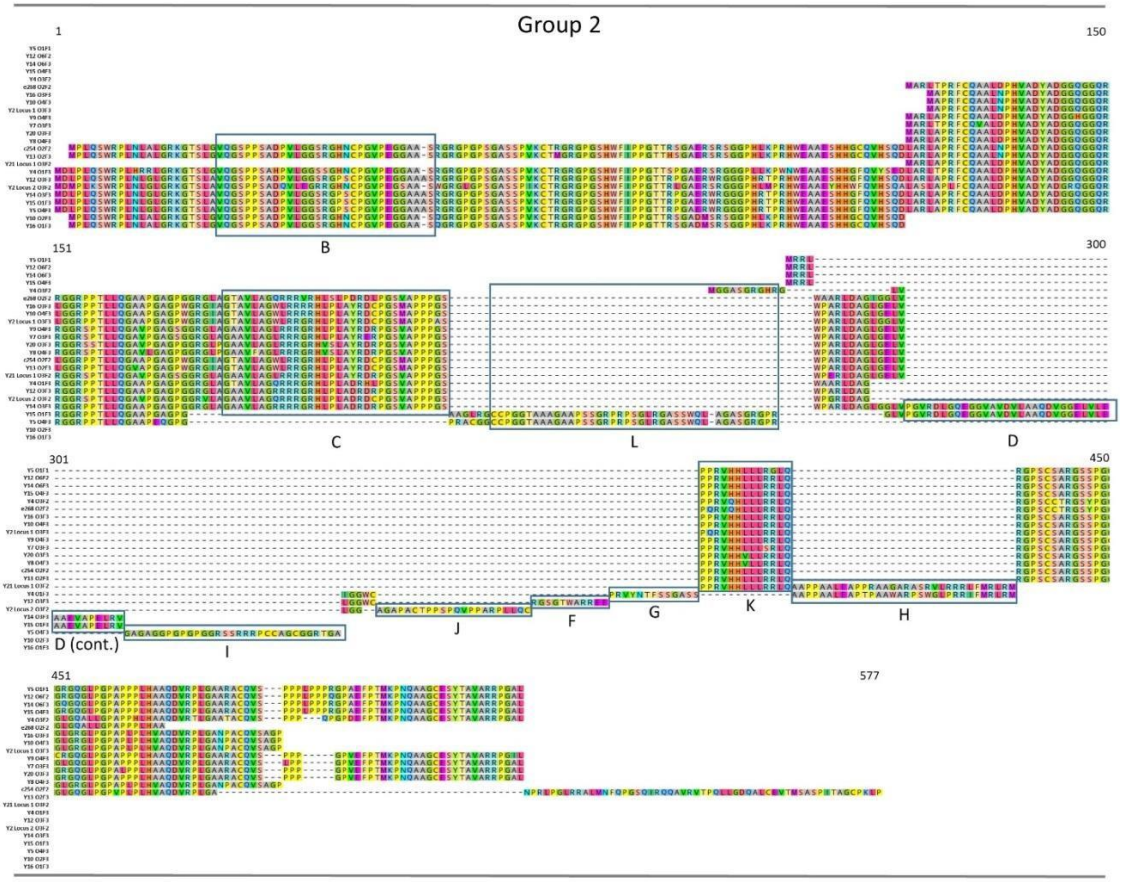


Figure A2.17. Protein alignments from the reverse open reading frames of *CCDC71L*, previously published *CCDC71LY* transcripts, and the Y-linked *CCDC71LY* loci identified in this study. Reverse open reading frames were divided into four groups based on similarity to enable cleaner alignments to greater facilitate comparisons. Shared amino acid segments between different groups are indicated in blocks and denoted by letters below each region.

CCDC71LY – Reverse ORFs



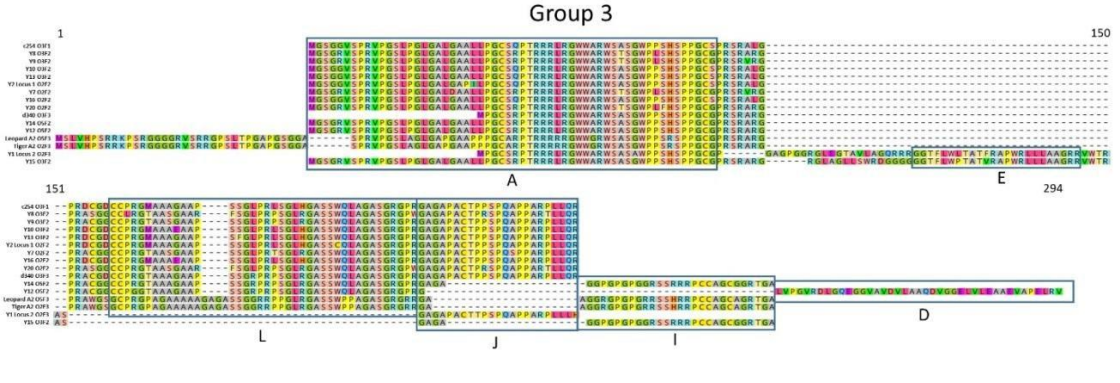


Figure A2.18. Protein alignments from the reverse open reading frames of *CCDC71L*, previously published *CCDC71LY* transcripts, and the Y-linked *CCDC71LY* loci identified in this study. Reverse open reading frames were divided into four groups based on similarity to enable cleaner alignments to greater facilitate comparisons. Shared amino acid segments between different groups are indicated in blocks and denoted by letters below each region.

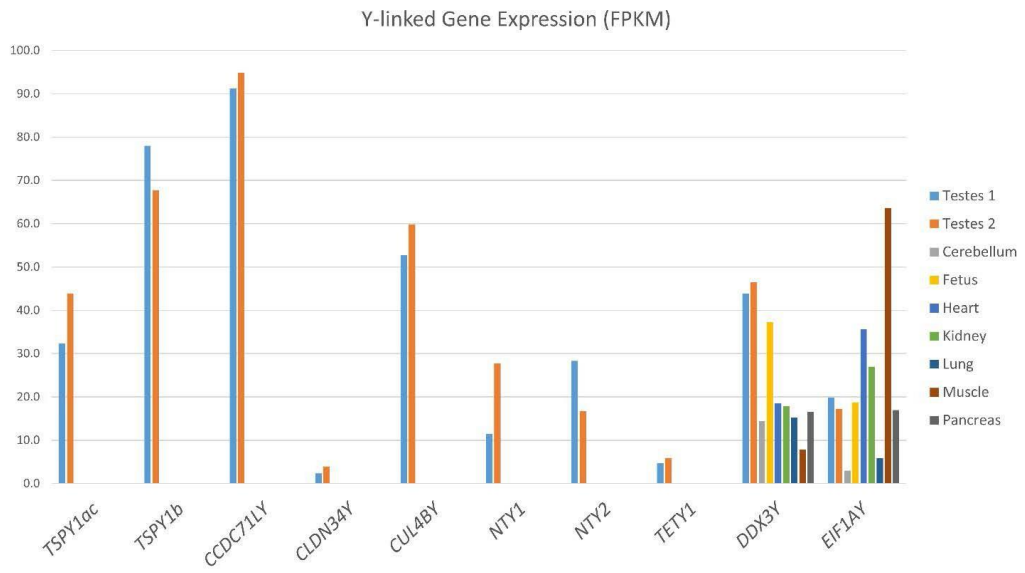


Figure A2.19. Tissue expression profile for ampliconic genes including two new genes identified in this study (*NTY1* and *NTY2*). Expression profiles for two single copy Y-linked genes (*DDX3Y* and *EIF1AY*) are also included as controls to contrast the testis-specific expression of the ampliconic genes.

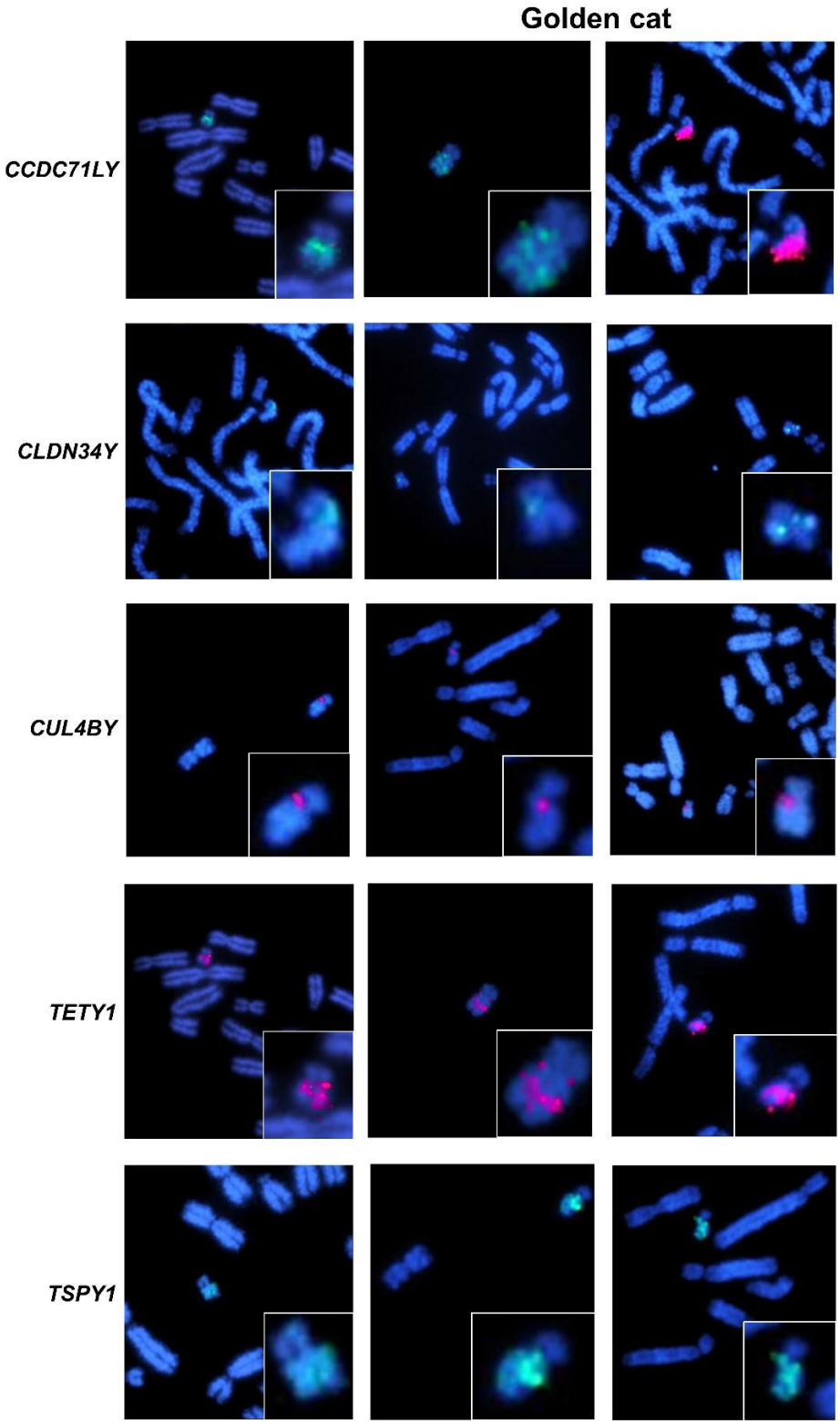


Figure A2.20. Additional images of fluorescent *in-situ* hybridization with probes designed for known ampliconic genes in metaphase preparations of Asian golden cat cells.

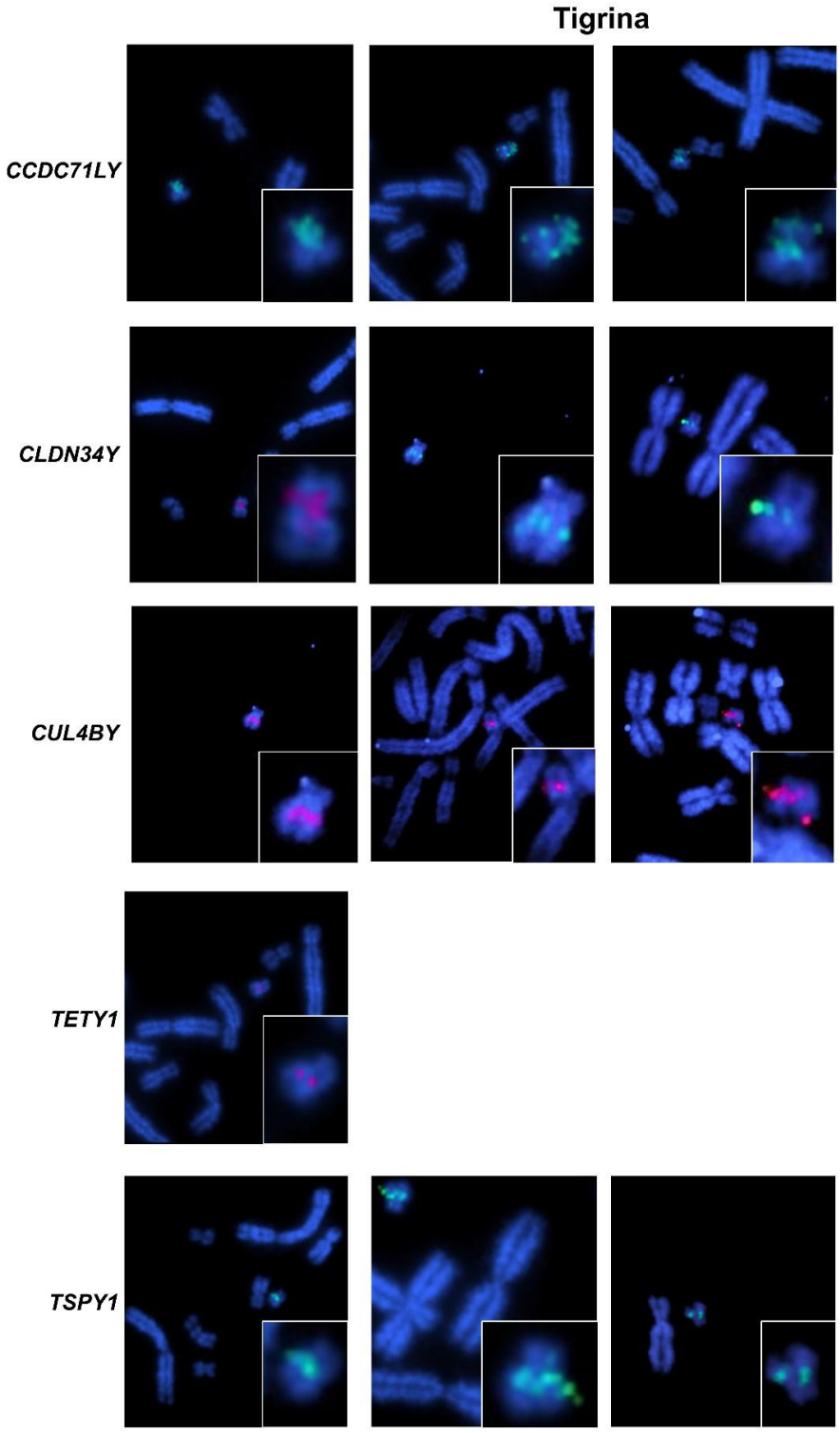


Figure A2.21. Additional images of fluorescent *in-situ* hybridization with probes designed for known ampliconic genes in metaphase preparations of tigrina cells.

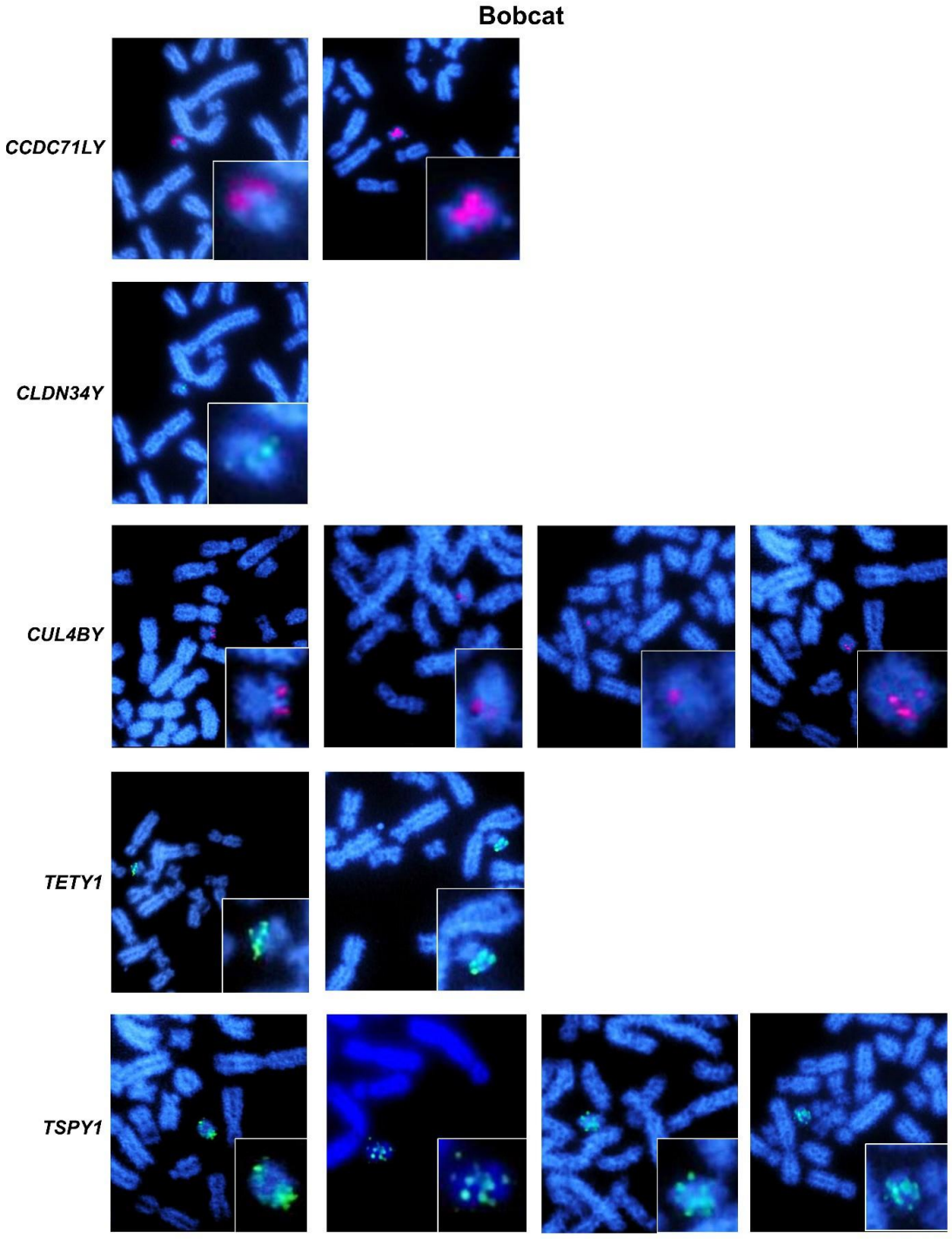


Figure A2.22. Additional images of fluorescent *in-situ* hybridization with probes designed for known ampliconic genes in metaphase preparations of bobcat cells.

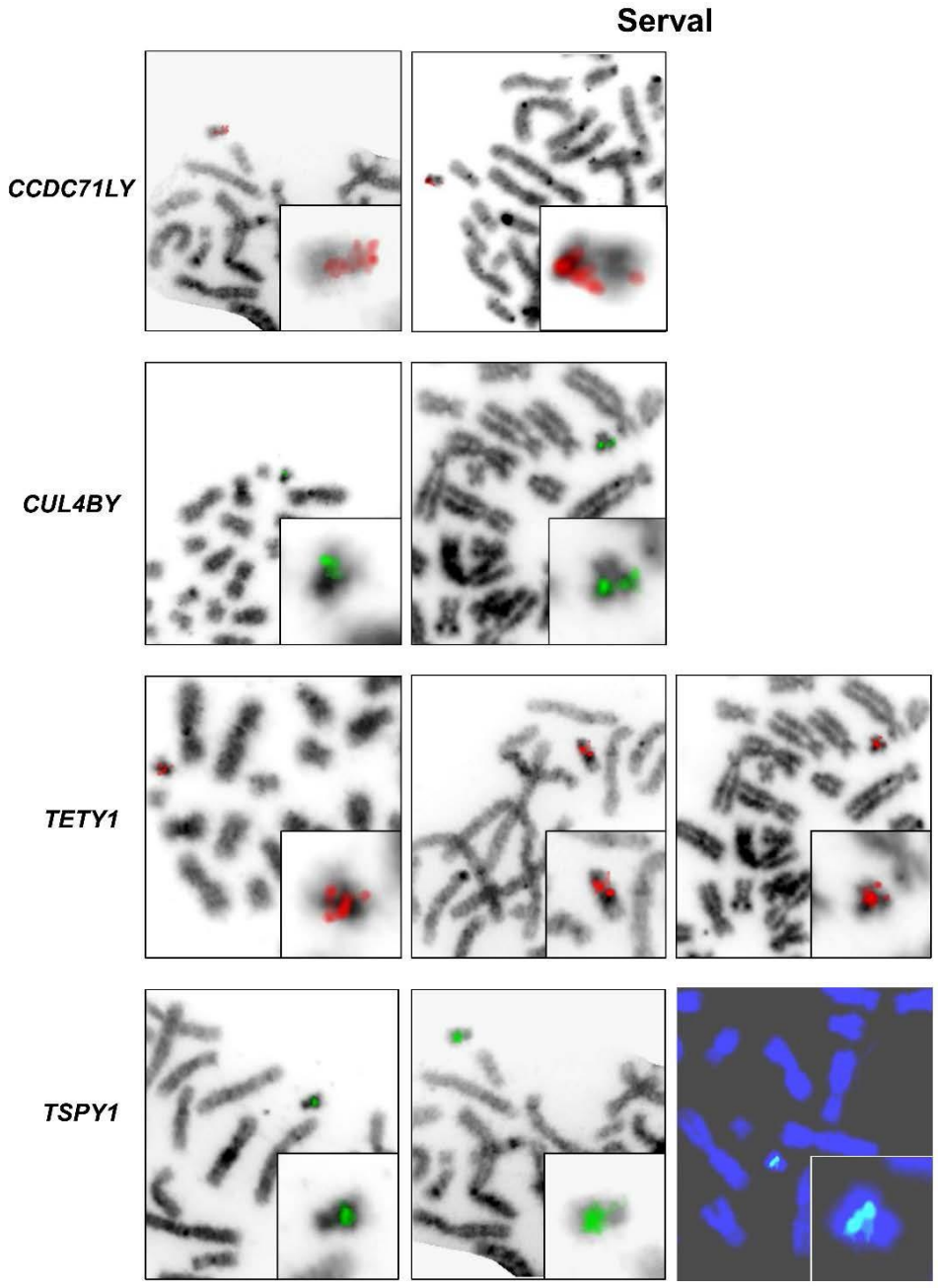


Figure A2.23. Additional images of fluorescent *in-situ* hybridization with probes designed for known ampliconic genes in metaphase preparations of serval cells.

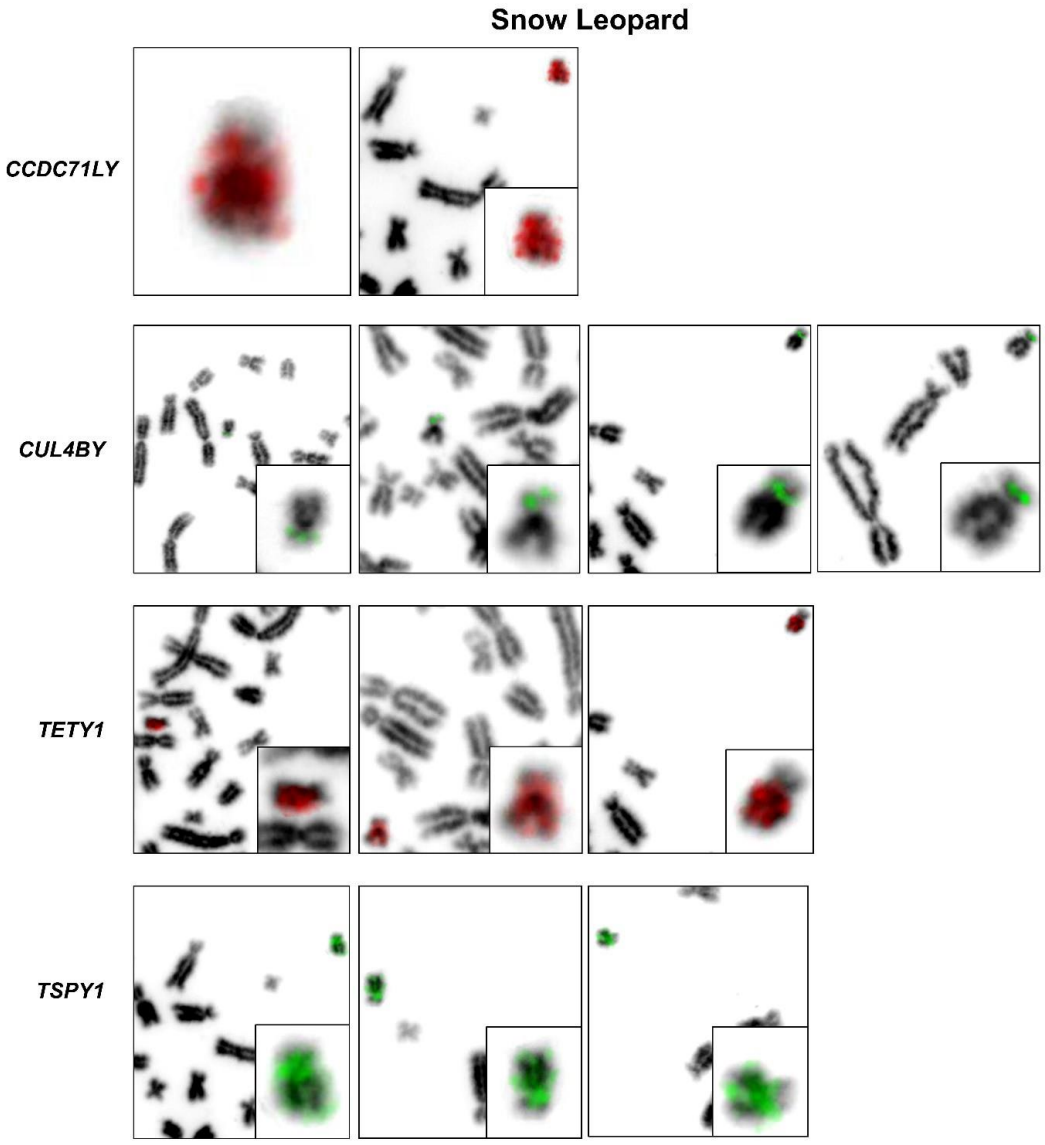
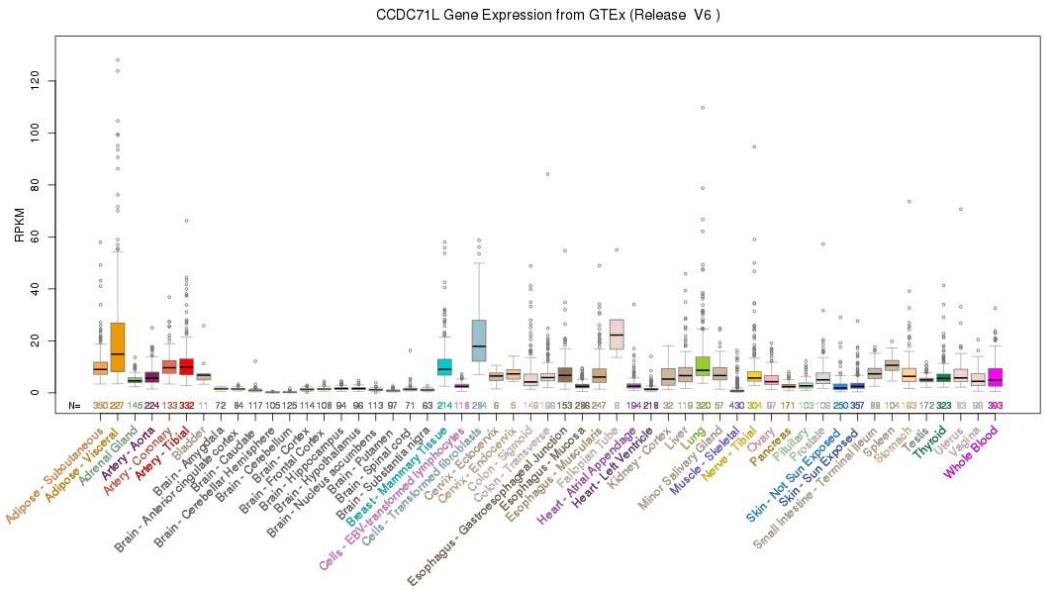


Figure A2.24. Additional images of fluorescent *in-situ* hybridization with probes designed for known ampliconic genes in metaphase preparations of snow leopard cells.


CCDC71L Gene Expression from GTEx (Release V6)
RPKM
0
20
40
60
80
100
120
N=
350 227 145 224 133 332 11 72 84 117 105 125 114 108 94 96 113 97 71 63 214 118 284 6 5 148 188 153 286 247 9 194 218 32 119 320 57 430 304 97 171 103 108 250 357 88 104 183 172 323 83 88 393
Adipose - Subcutaneous
Adipose - Visceral
Adrenal Gland
Artery - Aorta
Artery - Coronary
Artery - Tibial
Bladder
Brain - Amygdala
Brain - Anterior cingulate cortex
Brain - Caudate
Brain - Cerebellar Hemisphere
Brain - Cerebellum
Brain - Cortex
Brain - Frontal Cortex
Brain - Hippocampus
Brain - Hypothalamus
Brain - Nucleus accumbens
Brain - Putamen
Brain - Spinal cord
Brain - Substantia nigra
Breast - Mammary Tissue
Cells - EBV-transformed lymphocytes
Cells - Transformed fibroblasts
Cervix - Ectocervix
Cervix - Endocervix
Colon - Sigmoid
Colon - Transverse
Esophagus - Gastroesophageal Junction
Esophagus - Mucosa
Esophagus - Muscularis
Fallopian Tube
Heart - Atrial Appendage
Heart - Left Ventricle
Kidney - Cortex
Liver
Lung
Minor Salivary Gland
Muscle - Skeletal
Nerve - Tibial
Ovary
Pancreas
Pituitary
Prostate
Skin - Not Sun Exposed
Skin - Sun Exposed
Small Intestine - Terminal Ileum
Spleen
Stomach
Testis
Thyroid
Uterus
Vagina
Whole Blood

APPENDIX B

SUPPLEMENTAL TABLES

Table B2.1 Sex-biased ratios in hybrid felid crosses. Sources: http://www.ligercubs.com/liliger-cubs-in-russia, http://newsfeed.time.com/2013/06/20/meet-the-baby-liligers-born-in-a-russian-zoo, B. Davis: pers. comm.

Hybrid	**Number of litters**	**Number of Males**	**Number of females**	χ^2
Liliger (F1 - female backcrossed to lion)	5	3	13	0.012
Tiliger (F1 - female backcrossed to Tiger)	2	5	3	0.48
Liger (F1 - male lion crossed to female lion)	2	1	6	0.059

Table B2.2 Summary statistics of sequenced BAC clones.

Sequence	Sequencing Platform	Length (bps)	Scaffold Number	Contig Number
Y1	PacBio	141,646	1	1
Y2	Illumina MiSeq	133,556	1	2
Y3	Illumina MiSeq	126,499	1	4
Y4	PacBio	98,469	1	1
Y5	PacBio	124,712	1	1
Y6	Illumina MiSeq	106,795	2	3
Y7	Illumina MiSeq	88,899	1	4
Y8	Illumina MiSeq	112,421	1	3
Y9	Illumina MiSeq	131,753	1	1
Y10	Illumina MiSeq	96,072	1	1
Y11	Illumina MiSeq	88,866	1	3
Y12	PacBio	97,069	1	1
Y13	Illumina MiSeq	126,146	1	4
Y14	Illumina MiSeq	130,729	1	4
Y15	PacBio	101,612	1	1
Y16	Illumina MiSeq	123,022	1	6
Y17	Illumina MiSeq	110,508	1	8
Y18	PacBio	128,903	1	1
Y19	Illumina MiSeq	90,377	1	6
Y20	Illumina MiSeq	87,106	1	6
Y21	Illumina MiSeq	137,819	1	6
Total		*2,382,979*	*22*	*67*

Table B2.3 Comparison of assemblies from five clones that were sequenced using both the PacBio and Illumina MiSeq platforms.

Clone	Scaffold Name	Percent Identity (non-gapped sites)	%Recovered	Difference in scaffold length (MiSeq relative to PacBio)
529-O11	Y1	99.90%	97.83%	-3066 bp
532-F8	Y4	99.99%	99.98%	-18 bp
544-B22	Y12	99.96%	99.72%	-275 bp
554-L9	Y15	99.71%	98.86%*	+1196 bp
572-E6	Y18	99.77%	98.24%*	+2310 bp

* Percent recovered here represents the length of the PacBio relative to the MiSeq assembly.

Table B2.4. Genomic locations of endogenous retrovirus-like element enriched on the X and Y chromosomes.

Chromosome	Starting Coordinate	Ending Coordinate	Span
Y1	40140	48105	7965
Y1	88473	95519	7046
Y2	102631	109716	7085
Y3	740	8583	7843
Y4	8805	17488	8683
Y4	52999	60592	7593
Y6	26430	34039	7609
Y8	35262	43006	7744
Y9	13366	19837	6471
Y13	1190	8835	7645
Y13	40510	48269	7759
Y17	46941	54149	7208
Y17	86371	93382	7011
Y20	91	7155	7064
Y20	33402	41065	7663
Y21	45901	53529	7628
Y21	94003	102090	8087
A1	85482384	85487265	4881
A1	88767735	88771512	3777
A1	146217963	146225092	7129
A2	59796103	59801107	5004
A2	120954342	120962444	8102
A3	37693462	37700759	7297
B1	103081754	103089007	7253
B2	1333322	1340637	7315
B2	60931815	60938592	6777
B3	11512987	11520795	7808
B3	21786796	21791559	4763
B3	25250740	25252928	2188
B3	25252977	25258137	5160
B3	25828442	25830633	2191
B4	33331656	33338913	7257
B4	54669776	54677103	7327
C1	75900467	75907467	7000
C1	94793576	94798423	4847
C1	140257381	140264983	7602

Table B2.4 continued.

Chromosome	Starting Coordinate	Ending Coordinate	Span
C2	152905899	152911634	5735
C2	158853997	158861004	7007
D1	4778056	4783554	5498
D1	98176398	98183721	7323
D1	103712988	103720210	7222
D4	62680	68048	5368
D4	5919578	5924807	5229
D4	77321472	77328770	7298
D4	86719513	86726521	7008
E2	23093905	23101059	7154
E2	41370362	41377670	7308
E3	16334152	16339671	5519
E3	30047358	30054445	7087
X	5710319	5717595	7276
X	33173610	33179337	5727
X	51377431	51384676	7245
X	55120561	55128078	7517
X	58071347	58076132	4785
X	59495513	59502594	7081
X	64127327	64134605	7278
X	64410031	64413185	3154
X	66354061	66357521	3460
X	72173107	72180406	7299
X	72948316	72955396	7080
X	74298044	74302033	3989
X	78647525	78654565	7040
X	82043789	82051086	7297
X	83551453	83554842	3389
X	88534648	88547735	13087

Table B2.5. Estimated copy numbers of Y-linked ampliconic genes and autosomal *CCDC71L* across Felidae based on whole genome sequencing data. Estimates for *CCDC71L* were made based on reads mapping to the flanking regions of this gene in females. This was done because a gap in the assembly within the coordinates of this gene precluded accurate mapping of reads and to avoid overestimates of copy number due to Y-linked sequences incorrectly mapping to the autosome. Estimates for *TETY1* were based on the sum of reads mapping to both the Y-linked sequence and the autosomal progenitor, *TETYL*. A comparison of copy number estimates of *TETYL* in females and males confirm that the amplification of this region called by CNVnator were due to Y-derived reads incorrectly mapping to the autosomal copy.

Sample	Breed/Species	Sex	*TSPY ac*	*TSPY b*	5' *CCDC71L*	*CCDC71L*	3' *CCDC71L*	*CCDC71LY*	*CUL4BY*	*TETY1*	*TETY1L*	*TETY2*
SRR5040111	Domestic shorthair	M	76	40	NA	NA	NA	205	18	82	17	2
SRR5040117	Domestic shorthair	M	89	45	NA	NA	NA	250	22	103	22	2
SRR5040125	Domestic shorthair	M	81	31	NA	NA	NA	266	17	117	22	2
SRR5055389	Domestic shorthair	F	0	0	3	1	2	0	0	0	1	0
SRR5055404	Domestic shorthair	M	84	44	NA	NA	NA	213	21	90	16	2
SRR5038380	Oriental shorthair	M	108	40	NA	NA	NA	348	20	151	28	2
SRR5051114	Tennessee Rex	M	96	51	NA	NA	NA	238	21	107	22	2
SRR5051115	Devon Rex	F	0	0	2	1	2	0	0	0	1	0
SRR5055394	Domestic shorthair	M	88	49	NA	NA	NA	213	20	90	17	2
SRR5043257	Himalayan	M	91	50	NA	NA	NA	249	22	110	22	2
SRR5043264	Siamese	M	93	50	NA	NA	NA	248	22	112	22	2
SRR5055397	Birman	M	94	47	NA	NA	NA	243	21	103	18	2
SRR5055401	Burmese	M	85	37	NA	NA	NA	247	19	106	15	2
SRR066071-SRR066075	European wild cat	M	95	48	NA	NA	NA	270	20	111	26	NA
SRR2062538	Sand cat	M	123	51	NA	NA	NA	231	16	126	36	15

Table B2.5 continued.

Sample	Breed/Species	Sex	*TSPY1 ac*	*TSPY1 b*	5' *CCDC71L*	*CCDC71L*	3' *CCDC71L*	*CCDC71LY*	*CUL4BY*	*TETY 1*	*TETY1L*	*TETY 2*
SRR251181	Black footed cat	M	57	63	NA	NA	NA	93	38	32	18	21
SRR273751	Cheetah	M	75	101	NA	NA	NA	234	9	6	10	15
SRR273754	Cheetah	M	103	129	NA	NA	NA	213	5	4	7	2
SRR836361	Lion	M	51	52	NA	NA	NA	48	6	40	45	8
SRR836370	Lion	F	0	0	6	1	2	0	0	0	1	0
SRR304149	Leopard	F	0	0	6	1	3	0	0	0	1	0
SRR304222	Leopard	F	0	0	5	1	3	0	0	0	1	0
SRR1712667	Bengal Tiger	F	0	0	4	3	4	0	0	0	1	0
SRR836311	Siberian tiger	M	38	21	NA	NA	NA	104	6	13	27	2
SRR836354	Siberian tiger	M	11	0	NA	NA	NA	108	7	23	17	6
SRR836372	Snow leopard	F	0	0	7	1	4	0	0	0	1	0

Table B2.6. Copy number estimates based on quantitative real-time polymerase chain reactions conducted using tissues from male felids.

Species	***TSPY1ac***	***TSPY1b***	***CCDC71L***	***CCDC71LY***	***TETY2***
Domestic cat	96	88	1	124	3
Domestic cat	75	70	1	49	2
Chinese desert cat	130	96	1	128	7
Black footed cat	65	71	1	86	10
Jungle cat	71	48	1	110	82
Asian leopard cat	105	78	1	25	36
Asian leopard cat	166	66	1	48	40
Pallas cat	78	132	1	184	1
Iberian lynx	55	77	1	10	6
Margay	10	36	1	33	4
Pampas cat	35	64	1	36	3
Serval	52	105	1	43	65
Lion	140	161	26	88	22
Lion	140	167	27	147	20
Leopard	118	95	25	280	15
Jaguar	117	101	17	31	20
Jaguar	90	44	21	44	29
Tiger	54	43	15	12	12
Snow leopard	53	49	25	61	46
Clouded leopard	50	114	18	99	2

Table B2.7. NCBI SRA reads used for tissue expression analysis of Y-linked ampliconic genes.

SRA	**Description**	**Tissue**	**# of Bases**	**Disease/Noted Phenotype**
SRR3200449	Fcat18309	Lung	15.4G	Hydrocephaly
SRR3200450	Fcat17648	Fetus	25.2G	NA
SRR3200451	Fcat18309	Muscle	13.4G	Hydrocephaly
SRR3200462	Fcat11894	Testes	7.2Gb	Polycystic kidney disease
SRR3200469	Fcat18309	Pancreas	20.8G	Hydrocephaly
SRR3200471	Fcat18309	Heart	26.9G	Hydrocephaly
SRR3200473	Fcat11894	Kidney	11.8G	Polycystic kidney disease
SRR3218718	Fcat11894	Cerebellum	5.8G	Polycystic kidney disease
SRR1981105	FCA_4048	Testis	3.4G	NA

Table B2.8. Primers designed for qPCR based on multi-species alignments of known ampliconic Y-linked genes.

Locus	Forward	Reverse
DDX3Y	GAGATTCAGCGGAGGATTTG	CTGCTGCTGCTGTGACTACC
TSPY1ac	GGGTGTGAGATACAGGTGGC	CTTCTGGCTTGGCTTCCTCA
TSPY1b	CTACAGACGCAGGCATGACA	CCCTCTCAGTKCTCTCTCCT
CCDC71L	GCAGCCAGATTTTACTGCAA	AGAAGCTGCCAAGCAGAGAG
CCDC71LY	CCAAGTCCTCGGAGTAAACC	TTCATCCCTCCAGGCACAAC
CLDN34Y	GTGTGGGAGTTTAACAGC	GCCCTGCCCATACCAGATTT

Table B2.9. Samples downloaded from NCBI SRA database used for in silico (ITAL?) estimates of copy number variations of Y-linked ampliconic genes.

Sample	Breed/Species	Sex
SRR5040111	Domestic shorthair	Male
SRR5040117	Domestic shorthair	Male
SRR5040125	Domestic shorthair	Male
SRR5055389	Domestic shorthair	Female
SRR5055404	Domestic shorthair	Male
SRR5038380	Oriental shorthair	Male
SRR5051114	Tennessee Rex	Male
SRR5051115	Devon Rex	Female
SRR5055394	Domestic shorthair	Male
SRR5043257	Himalayan	Male
SRR5043264	Siamese	Male
SRR5055397	Birman	Male
SRR5055401	Burmese	Male
SRR066071-SRR066075	European wild cat	Male
SRR2062538	Sand cat	Male
SRR2511861	Black footed cat	Male
SRR2737521	Cheetah	Male
SRR2737544	Cheetah	Male
SRR836361	Lion	Male
SRR836370	Lion	Female
SRR3041429	Amur leopard	Female
SRR3042212	Amur leopard	Female
SRR1712667	Bengal Tiger	Female
SRR836311	Siberian tiger	Male
SRR836354	Siberian tiger	Male
SRR836372	Snow leopard	Female

Table B3.1. BAC clones from the Felis catus female BAC library FSCC from Amplicon Express sequenced and assembled for this study.

Clone	Assembly Length	X Start	X End	BLAST Length	Orientation	Gap	Gap Spanned	Identity	Query Coverage
FCAB-33C13	238,715	20,834,159	21,072,609	238,450	reverse	No	NA	99.9+%	99.90%
FCAB-252P24	169,583	22,100,405	22,272,041	171,636	forward	No	NA	99.9+%	99.60%
FCAB-119B21	183,911	22,410,045	22,593,683	183,638	reverse	No	NA	99.9+%	99.80%
FCAB-249D18	175,324	23,607,503	23,784,075	176,572	forward	No	NA	99.9+%	99.9+%
FCAB-204B24	179,016	23,943,803	24,122,813	179,010	reverse	No	NA	99.9+%	99.9+%
FCAB-8C17	134,432	32,334,177	32,463,016	128,839	reverse	No	NA	99.9+%	94.90%
FCAB-185N2	127,836	32,442,450	32,546,221	103,771	forward	32,517,052-32,517,152	Yes	99.9+%	74.90%
FCAB-142K3	111,769	32,486,718	32,602,742	116,024	forward	32,517,052-32,517,152	Yes	99.9+%	99.40%
FCAB-8C17-185N2	241,687	32,334,177	32,546,221	212,044	forward	32,517,052-32,517,152	Yes	99.9+%	83.90%
FCAB-150K12	184,107	33,158,996	33,348,055	189,059	reverse	33,298,646-33,299,144	Yes	99.9+%	91.20%
FCAB-208C24	160,042	Ends in gap	33,863,224	NA	forward	33,652,424-33,708,759	No	99.80%	88.50%
FCAB-121A12	257,950	41,948,456	42,209,759	261,303	reverse	No	NA	99.80%	99.9+%
FCAB-95E10	239,858	42,282,090	42,529,989	247,899	forward	No	NA	99.90%	99.90%
FCAB-100P9	237,719	42,502,450	42,736,322	233,872	forward	No	NA	99.9+%	96.60%
FCAB-95E10-100P9	449,985	42,282,090	42,736,322	454,232	forward	No	NA	99.9+%	98.20%
FCAB-173P5	244,847	43,390,319	43,634,864	244,545	reverse	No	NA	99.9+%	99.70%
FCAB-201P4	175,838	44,977,504	45,154,735	177,231	reverse	No	NA	99.9+%	99.80%
FCAB-259J16	209,572	45,251,849	45,461,127	209,278	forward	No	NA	99.9+%	99.80%
FCAB-204F6	211,725	Ends in gap	46,197,369	NA	reverse	46,047,039-46,047,138	No	99.9+%	70.90%
FCAB-31J21	205,850	46,642,669	Ends in gap	NA	reverse	46,806,000-46,806,100	No	99.9+%	78.30%
FCAB-174L20	131,839	46,793,384	46,837,270	43,886	reverse	46,806,000-46,806,099	Yes	99.9+%	33.20%
FCAB-31J21-174L20	280,356	46,642,669	46,837,270	194,601	forward	46,806,000-46,806,100	Yes	99.9+%	68.60%

Table B3.1. Continued

Clone	Assembly Length	X Start	X End	BLAST Length	Orientation	Gap	Gap Spanned	Identity	Query Coverage
FCAB-238C17	169,848	47,525,890	47,696,251	170,361	forward	No	NA	99.9+%	99.9+%
FCAB-120M10	163,279	47,588,289	47,752,040	163,751	forward	No	NA	99.9+%	99.40%
FCAB-287E2	202,415	47,598,545	47,801,430	202,885	forward	No	NA	99.9+%	99.50%
FCAB-275M22	138,636	47,629,840	47,768,727	138,887	reverse	No	NA	99.9+%	99.20%
FCAB-39C5	188,732	47,786,959	47,978,054	191,095	reverse	No	NA	99.9+%	99.9+%
FCAB-238C17-120M10-287E2-275M22-39C5	449,261	47,525,890	47,978,054	452,164	forward	No	NA	99.9+%	99.70%
FCAB-83L9	151,285	48,047,088	48,199,192	152,104	reverse	No	NA	99.9+%	99.9+%
FCAB-117E18	186,808	Ends in gap	48,630,108	NA	reverse	48,422,751; 48,515,913	48,515,913 spanned	99.90%	87.30%
FCAB-205P3	233,420	Unclear	Unclear	NA	reverse	49476788; 49510742	Unclear	99.80%	56.20%
FCAB-309N18	133,097	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear
FCAB-86B10	99,124	49,615,208	49,714,555	99,347	reverse	No	NA	99.9+%	98.30%
FCAB-237G9	157,486	49,650,370	49,808,094	157,724	reverse	No	NA	99.9+%	99.40%
FCAB-141O2	164,325	49,743,731	49,923,230	179,499	reverse	No	NA	99.9+%	99.70%
FCAB-127P13	158,243	49,865,345	50,027,913	162,568	reverse	No	NA	99.9+%	99.70%
FCAB-205P3-309N18-86B10-237G9-141O2-127P13	629,752	49,527,913	50,027,913	500,000	forward	No	NA	99.9+%	73.50%
FCAB-17I12	139,800	50,253,823	50,355,933	102,110	forward	50,326,189-50,326,288	Yes	99.9+%	52.20%
FCAB-179D2	155,554	Ends in gap	52,652,819		forward	50,521,358-52,521,357	No	99.9+%	91.80%
FCAB-137K7	189,068	53,342,260	53,576,912	234,652	reverse	No	NA	99.9+%	99.9+%
FCAB-271N13	260,836	57,134,058	57,395,674	261,616	forward	No	NA	99.9+%	99.70%
FCAB-256L9	175,642	61,911,666	62,108,014	196,348	reverse	61,956,860	Yes	99.9+%	99.9+%
FCAB-100A3	137,026	62,466,700	62,605,414	138,714	forward	No	NA	99.9+%	99.9+%
FCAB-261A2	157,020	62,714,864	62,871,983	157,119	forward	No	NA	99.9+%	99.9+%
FCAB-182J5	208,434	63,836,592	64,057,428	220,836	forward	No	NA	99.9+%	99.80%
FCAB-278C20	157,811	63,938,560	64,103,141	164,581	forward	No	NA	99.80%	99.80%
FCAB-182J5-278C20	253,589	63,836,592	64,103,141	266,549	forward	No	NA	99.90%	99.80%

Table B3.1. Continued

Clone	Assembly Length	X Start	X End	BLAST Length	Orientation	Gap	Gap Spanned	Identity	Query Coverage
FCAB-158E6	187,851	64,694,492	64,883,566	189,074	reverse	No	NA	99.9+%	99.80%
FCAB-226I7	221,997	64,979,879	65,206,789	226,910	reverse	No	NA	99.9+%	99.40%
FCAB-314I22	233,015	65,774,740	66,008,622	233,882	forward	No	NA	99.9+%	99.20%
FCAB-184D11	183,500	74,939,604	Ends in gap	NA	reverse	75,084,129-75,084,228	No	99.90%	65.60%
FCAB-44I3	92,899	75,035,654	Ends in gap	NA	reverse	75,084,129-75,084,228	No	99.60%	26.30%
FCAB-184D11-44I3	188,891	74,939,604	Ends in gap	NA	forward	75,084,129-75,084,228	No	99.90%	63.70%
FCAB-205N19	218,342	77,597,851	77,816,551	218,700	reverse	No	NA	99.9+%	99.90%
FCAB-154F16	230,586	80,448,478	80,678,552	230,074	reverse	No	NA	99.70%	99.9+%
FCAB-326J11	154,759	Ends in gap	84,961,267	NA	reverse	84,837,303-84,837,801	No	99.60%	49.30%
FCAB-192L23	119,007	85,807,622	Ends in gap	NA	reverse	85,921,508	No	99.9+%	95.70%
FCAB-265O20	133,961	86,051,873	86,187,045	135,172	reverse	No	NA	99.9+%	99.80%
FCAB-151N4	236,889	86,070,152	86,308,140	237,988	forward	No	NA	99.9+%	99.90%
FCAB-265020-151N4	253,749	86,051,873	86,308,140	256,267	forward	No	NA	99.9+%	99.90%
FCAB-212B23	122,276	88,591,659	88,714,980	123,321	reverse	No	NA	99.9+%	95.70%
FCAB-52K13	112,352	88,817,443	88,929,875	112,432	reverse	No	NA	99.9+%	99.30%
FCAB-92N14	106,781	88,823,027	88,929,883	106,856	reverse	No	NA	99.9+%	99.20%
FCAB-137M13	105,310	88,824,490	88,929,875	105,385	reverse	No	NA	99.9+%	99.20%
FCAB-185B10	139,897	88,824,513	88,964,508	139,995	reverse	No	NA	99.9+%	99.40%
FCAB-52K13-92N14-137M13-185B10	146,961	88,817,443	88,964,508	147,065	forward	No	NA	99.9+%	99.40%
FCAB-319N21	157,163	95,638,746	95,799,266	160,520	forward	No	NA	99.90%	99.80%
FCAB-208B12	192,591	96,459,107	96,651,711	192,604	reverse	No	NA	99.9+%	99.9+%
FCAB-206I16	184,413	97,552,011	97,736,027	184,016	forward	No	NA	99.9+%	99.10%
FCAB-88D1	207,808	100,686,264	100,893,798	207,534	forward	No	NA	99.9+%	99.60%
FCAB-165E15	244,297	100,835,916	101,085,257	249,341	forward	No	NA	99.9+%	99.9+%
FCAB-88D1-165E15	364,221	100,686,264	101,085,257	398,993	forward	No	NA	99.9+%	99.80%

Table B3.1. Continued

Clone	Assembly Length	X Start	X End	BLAST Length	Orientation	Gap	Gap Spanned	Identity	Query Coverage
FCAB-208P17	109,731	103,282,229	103,392,918	110,689	forward	No	NA	99.9+%	99.9+%
FCAB-77K22	195,219	103,386,042	103,583,587	197,545	reverse	No	NA	99.9+%	99.9+%
FCAB-208P17-77K22	298,065	103,282,229	103,583,587	301,358	forward	No	NA	99.9+%	99.9+%
FCAB-89F20	157,935	113,531,770	Ends in gap	NA	forward	113,681,637-113,789,782	No	99.9+%	94.90%
FCAB-2D2	135,365	Ends in gap	113,905,251	NA	forward	113,681,637-113,789,782	No	99.9+%	85.10%
FCAB-296I24	269,552	122,804,294	123,076,639	272,345	reverse	No	NA	99.9+%	99.90%
FCAB-68L12	192,460	123,030,652	Ends in gap	NA	reverse	123,218,969	No	99.9+%	97.30%
FCAB-77K16	203,921	123,142,491	123,346,723	204,232	reverse	123,218,969	Yes	99.9+%	89.10%
FCAB-157F17	138,162	123,208,238	123,346,723	138,485	reverse	123,218,969	Yes	99.9+%	84.00%
FCAB-296I24-68L12-77K16-157F17	539,319	122,804,294	123,346,723	542,429	forward	123,218,969	Yes	99.9+%	95.60%
FCAB-70K1	224,711	123,372,052	123,596,619	224,567	forward	No	NA	99.9+%	99.80%
FCAB-146O7	171,640	125,176,746	125,454,965	278,219	reverse	125,178,926	No	99.80%	93.30%
FCAB-286M16	155,991	127,154,492	127,310,399	155,907	reverse	No	NA	99.9+%	99.90%
FCAB-60D9	200,388	127,912,736	Ends in gap	NA	forward	128,222,943; 128,343,109	Unclear	99.9+%	65.50%
FCAB-197A7	252,015	127,912,736	Ends in gap	NA	forward	128,222,943; 128,343,110	Unclear	99.80%	51.90%
FCAB-201F19	269,191	128,240,391	Ends in gap	NA	reverse	128,222,943; 128,343,110	Unclear	99.80%	37.90%
FCAB-174G3	136,721	Ends in gap	Ends in gap	NA	forward	128,222,943; 128,343,110	Unclear	99.80%	87.00%
FCAB-278F23	121,693	128,240,354	Ends in gap	NA	forward	128,343,110	Unclear	99.80%	83.90%
FCAB-60D9-197A7-201F19-174G3-278F23	287,046	127,912,736	128,178,239	265,503	forward	128,222,943; 128,343,110	Perhaps	99.90%	46.70%
FCAB-259D11	162,325	129,272,280	Ends in gap	NA	forward	129,435,018-129,435,117	Unclear	99.9+%	93.50%
FCAB-331E2	215,372	Ends in gap	129,505,645	NA	reverse	129,435,018-129,435,117	Unclear	99.9+%	16.90%
FCAB-97M22	219,043	130,047,672	130,266,800	219,128	forward	No	NA	99.9+%	99.9+%
FCAB-253K23	156,266	130,278,199	130,434,507	156,308	reverse	No	NA	100.00%	99.9+%
FCAB-111F22	176,814	130,302,694	130,479,169	176,475	forward	No	NA	99.9+%	99.70%
FCAB-253K23-111F22	201,251	130,278,199	130,479,169	200,970	forward	No	NA	99.9+%	99.80%

Table B3.2. Read mapping statistics from FelCat9.0 and our improved (9.1) assembly with incorporated sequenced BAC assemblies.

Clones	9.0 Error Rate	9.0 Avg Qual	9.0 Avg Insert Size	9.0 Insert Size S.D.	9.0 Reads Mapped	9.1 Error Rate	9.1 Avg Qual	9.1 Avg Insert Size	9.1 Insert Size S.D.	9.1 Reads Mapped
FCAB-33C13	1.8193E-03	35.0	333.5	107.6	92,188	1.6568E-03	35.0	333.5	107.7	92,315
FCAB-252P24	1.9295E-03	35.1	335.3	110.5	67,389	1.7540E-03	35.2	335.1	110.1	71,132
FCAB-119B21	1.6856E-03	35.2	335.5	111.1	75,280	1.6685E-03	35.2	335.6	111.1	75,331
FCAB-249D18	1.8618E-03	35.2	335.5	111.0	71,835	1.6182E-03	35.2	335.2	110.6	72,158
FCAB-204B24	1.7242E-03	35.2	337.9	112.1	74,542	1.6343E-03	35.2	337.9	112.1	74,490
FCAB-8C17-185N2	3.5115E-03	35.1	356.0	189.3	135,297	2.3024E-03	35.1	335.1	108.2	133,834
FCAB-150K12	2.4366E-03	35.1	335.5	112.8	88,278	1.9770E-03	35.1	335.0	108.9	84,636
FCAB-208C24	1.8672E-03	35.1	335.9	110.3	58,345	1.8566E-03	35.1	334.7	108.9	74,847
FCAB-121A12	2.8103E-03	34.9	331.0	105.4	96,795	2.6371E-03	34.9	330.8	105.1	97,036
FCAB-95E10-100P9	2.9864E-03	35.0	333.0	107.4	171,086	2.6253E-03	35.0	333.1	107.1	171,428
FCAB-173P5	3.2700E-03	35.1	334.5	109.5	98,491	3.0649E-03	35.1	333.9	108.7	100,186
FCAB-201P4	2.0760E-03	35.1	336.1	111.1	73,123	1.9208E-03	35.1	335.8	110.2	74,538
FCAB-259J16	1.9512E-03	35.2	336.2	111.4	86,583	2.2241E-03	35.2	335.4	110.7	89,774
FCAB-204F6	2.0916E-03	35.0	333.9	107.3	56,422	2.9326E-03	35.0	333.4	107.3	96,211
FCAB-31J21-174L20	2.4328E-03	35.1	335.2	109.4	82,093	2.0211E-03	35.1	335.2	109.1	121,111
FCAB-238C17-120M10-287E2-275M22-39C5	2.2574E-03	35.1	335.5	109.8	184,433	1.9854E-03	35.1	335.4	109.3	187,548
FCAB-83L9	1.8907E-03	35.1	336.8	111.2	61,900	2.1749E-03	35.1	336.3	110.8	62,276
FCAB-117E18	1.9564E-03	35.2	337.6	111.8	68,984	1.8867E-03	35.2	337.2	111.5	86,073
FCAB-205P3-309N18-86B10-237G9-141O2-127P13	2.0631E-03	35.1	336.4	111.1	196,496	1.9038E-03	35.1	335.6	109.9	243,553
FCAB-17I12	1.9602E-03	35.1	335.5	109.8	39,398	1.9128E-03	35.1	335.4	110.4	57,913
FCAB-179D2	1.9406E-03	35.1	336.7	111.2	46,226	1.9455E-03	35.1	336.6	111.9	58,180
FCAB-137K7	1.9869E-03	35.2	336.5	111.6	96,150	1.7998E-03	35.2	336.1	110.7	78,255
FCAB-271N13	2.1662E-03	35.2	336.1	111.3	106,478	1.8436E-03	35.2	335.6	110.6	109,355
FCAB-256L9	2.2612E-03	35.1	335.2	109.7	71,273	2.1220E-03	35.1	335.6	109.5	73,632
FCAB-100A3	2.3962E-03	35.1	334.8	110.4	56,546	2.2135E-03	35.1	334.3	109.8	56,570
FCAB-261A2	2.3328E-03	35.1	336.1	110.4	65,990	2.2121E-03	35.1	336.2	110.4	66,236
FCAB-182J5-278C20	2.3969E-03	35.1	335.6	111.3	106,387	2.2203E-03	35.1	334.9	110.0	108,634
FCAB-158E6	2.2072E-03	35.2	337.1	111.5	77,586	1.9820E-03	35.2	336.9	111.1	79,251
FCAB-226I7	2.2989E-03	35.1	336.9	110.5	94,418	2.1624E-03	35.1	336.4	110.0	93,645
FCAB-314I22	2.3665E-03	35.2	336.1	111.3	94,022	2.1647E-03	35.2	335.8	110.7	97,302
FCAB-184D11-44I3	2.2703E-03	35.1	336.2	110.6	52,438	2.1639E-03	35.2	335.6	110.3	94,196
FCAB-205N19	2.9299E-03	35.2	335.7	110.4	89,540	2.7034E-03	35.2	335.5	110.0	90,237

Table B3.2. Continued

Clones	9.0 Error Rate	9.0 Avg Qual	9.0 Avg Insert Size	9.0 Insert Size S.D.	9.0 Reads Mapped	9.1 Error Rate	9.1 Avg Qual	9.1 Avg Insert Size	9.1 Insert Size S.D.	9.1 Reads Mapped
FCAB-154F16	1.9037E-03	35.2	336.2	111.2	94,831	1.8240E-03	35.2	336.3	111.2	95,319
FCAB-326J11	2.4948E-03	35.1	334.5	108.7	63,259	2.4268E-03	35.1	335.1	109.0	84,513
FCAB-192L23	2.2486E-03	35.1	335.5	109.6	89,636	1.9474E-03	35.1	337.0	111.2	52,031
FCAB-265020-151N4	2.3441E-03	35.1	336.6	111.2	102,681	2.2110E-03	35.1	336.2	110.5	51,886
FCAB-212B23	1.9431E-03	35.2	337.0	111.2	48,784	3.1609E-03	35.2	336.8	111.6	58,184
FCAB-52K13-92N14-137M13-185B10	1.8091E-03	35.2	335.9	110.1	55,988	1.7446E-03	35.2	336.1	109.9	63,138
FCAB-319N21	3.7586E-03	35.0	333.6	108.4	61,192	3.5026E-03	35.0	333.1	107.8	61,310
FCAB-208B12	1.8749E-03	35.0	333.4	108.5	75,935	1.8937E-03	35.0	333.1	108.1	77,536
FCAB-206I16	1.9632E-03	35.0	332.2	107.2	69,535	1.8891E-03	35.0	332.2	106.9	70,315
FCAB-88D1-165E15	2.0578E-03	34.9	331.9	106.3	147,697	1.9101E-03	34.9	331.9	106.4	148,301
FCAB-208P17-77K22	1.8263E-03	35.0	334.6	108.7	119,245	1.7150E-03	35.0	334.5	108.7	119,275
FCAB-89F20	2.6197E-03	34.9	331.7	106.9	57,927	2.3515E-03	34.9	331.4	106.2	61,175
FCAB-2D2	2.3740E-03	34.9	331.1	105.9	40,557	2.4096E-03	34.9	331.7	105.9	59,139
FCAB-296I24-68L12-77K16-157F17	1.8143E-03	35.2	336.7	111.6	212,705	1.6293E-03	35.2	336.6	111.3	221,409
FCAB-70K1	1.8378E-03	35.1	335.8	109.9	89,608	1.7402E-03	35.1	335.9	110.0	89,755
FCAB-146O7	3.2880E-03	35.1	594.2	1374.0	66,625	3.6481E-03	34.9	332.6	107.8	110,293
FCAB-286M16	1.9490E-03	35.1	336.7	110.7	65,185	1.8634E-03	35.1	336.4	110.2	65,892
FCAB-60D9-197A7-201F19-174G3-278F23	2.0846E-03	35.0	333.3	107.3	93,109	2.4747E-03	34.9	332.0	105.7	92,169
FCAB-259D11	3.1823E-03	34.6	336.9	149.9	55,960	2.0544E-03	34.7	328.1	101.5	53,786
FCAB-331E2	2.3012E-03	34.9	332.3	108.1	26,406	2.3997E-03	34.7	329.2	104.2	85,834
FCAB-97M22	2.0754E-03	35.2	337.1	111.1	92,442	1.9069E-03	35.2	337.1	110.9	93,595
FCAB-253K23-111F22	1.8525E-03	35.2	335.5	110.1	78,296	1.7885E-03	35.2	335.0	109.2	79,367

Table B3.3. Manual alignment of human, domestic cat, pig, and mouse X chromosomes based on conserved gene content. Genes within ampliconic regions are italicized, underlined, and in bold. Ampliconic regions without genes are indicated in black with white text. Novel genes in identified in this study within the cat are italicized and named according to the clone in which they were identified. When these transcripts were not found in the FelCat9.0 assembly, the coordinates which were replaced in our updated assembly are provided and indicated with an asterisk.

Gene Name	Human Start	Human End	Cat Start	Cat End	Pig Start	Pig End	Mouse Start	Mouse End
PLCXD1	276,324	303,356	57,166	82,324	Unplaced	Unplaced	-	-
LOC101098289	-	-	85,641	98,815	-	-	-	-
GTPBP6	304,750	318,819	119,686	131,954	Unplaced	Unplaced	-	-
LOC102901245	-	-	134,637	198,935	-	-	-	-
LOC111558781	-	-	199,003	207,366	-	-	-	-
PPP2R3B	333,933	386,955	207,276	267,874	Unplaced	Unplaced	-	-
SHOX	624,344	659,411	530,158	539,931	Unplaced	Unplaced	-	-
CRLF2	1,190,437	1,212,762	912,887	936,497	Unplaced	Unplaced	-	-
CSF2RA	1,268,800	1,325,097	975,880	994,473	Unplaced	Unplaced	-	-
LOC102901718	-	-	1,000,391	1,023,639	-	-	-	-
IL3RA	1,336,575	1,382,689	1,030,029	1,055,668	Unplaced	Unplaced	-	-
SLC25A6	1,386,152	1,392,146	1,055,595	1,059,018	Unplaced	Unplaced	-	-
LOC105373102	1,392,349	1,396,881	-	-	-	-	-	-
ASMTL	1,403,139	1,453,794	1,064,129	1,086,806	Unplaced	Unplaced	-	-
P2RY8	1,462,572	1,537,488	1,094,182	1,138,047	Unplaced	Unplaced	-	-
LOC109496913	-	-	1,135,898	1,139,304	-	-	-	-
LOC111558770	-	-	1,156,778	1,159,735	-	-	-	-
AKAP17A	1,591,593	1,602,520	1,156,797	1,171,088	Unplaced	Unplaced	-	-
ASMT	1,595,455	1,643,081	1,182,975	1,225,823	Unplaced	Unplaced	170,672,644	170,678,054
DHRSX	2,219,506	2,500,974	1,279,685	1,542,722	-	-	-	-
ZBED1	2,486,414	2,500,967	1,532,444	1,546,462	Unplaced	Unplaced	-	-
CD99	2,691,133	2,741,309	1,633,819	1,665,788	-	-	-	-

Table B3.3. Continued

Gene Name	Human Start	Human End	Cat Start	Cat End	Pig Start	Pig End	Mouse Start	Mouse End
XG	2,751,108	2,816,500	1,666,419	1,704,985	11,104	31,716	-	-
GYG2	2,828,712	2,882,820	1,719,668	1,751,867	52,310	89,412	-	-
ARSD	2,903,970	2,929,375	1,755,393	1,782,760	107,992	125,985	-	-
ARSE	2,934,632	2,968,310	1,782,867	1,800,955	131,769	156,083	-	-
ARSH	3,006,613	3,033,385	1,809,229	1,831,713	-	-	-	-
ARSF	3,040,234	3,112,729	1,841,434	1,872,523	184,977	220,103	-	-
LOC110257789	-	-	-	-	418,073	433,468	-	-
MXRA5	3,308,565	3,347,098	1,985,406	2,015,594	486,396	517,329	-	-
Cldn34b3	-	-	-	-	-	-	76,264,306	76,267,673
Cldn34b4	-	-	-	-	-	-	76,385,323	76,397,979
Cldn34d	-	-	-	-	-	-	76,582,610	76,602,924
PRKX	3,604,343	3,713,634	2,162,556	2,238,207	657,421	753,739	77,760,972	77,796,222
OBP	-	-	2,972,471	3,088,371	2,012,066	2,019,074	77,837,898	77,853,500
Gm14744 (OBP duplicate)	-	-	-	-	-	-	77,864,758	77,870,033
5430402E10Rik (Obp1-like)	-	-	-	-	-	-	77,919,786	77,925,062
Obp1a	-	-	-	-	-	-	78,085,505	78,091,374
Gm5938	-	-	-	-	-	-	78,125,467	78,130,403
Obp1b	-	-	-	-	-	-	78,186,808	78,192,242
Gm14743	-	-	-	-	-	-	78,240,462	78,245,921
4930480E11Rik	-	-	-	-	-	-	78,369,643	78,371,128
LOC110257836	-	-	-	-	2,127,385	2,145,225	-	-
Human Ampliconic	***3,821,643***	***3,938,401***	-	-	-	-	-	-
NLGN4X	5,890,026	6,228,882	3,662,083	3,986,209	2,701,397	3,103,824	-	-
VCX3A	6,533,618	6,535,118	-	-	-	-	-	-
PUDP	6,768,840	7,148,190	4,213,632	4,703,425	3,577,487	3,876,142	-	-

Table B3.3. Continued

Gene Name	Human Start	Human End	Cat Start	Cat End	Pig Start	Pig End	Mouse Start	Mouse End
STS	7,147,252	7,354,643	4,702,017	4,862,259	3,926,027	4,076,900	-	-
VCX	7,842,262	7,844,143	-	-	-	-	-	-
PNPLA4	7,898,763	7,928,587	5,344,916	5,379,562	4,540,846	4,623,831	-	-
VCX2	8,169,944	8,171,267	-	-	-	-	-	-
VCX3B	8,464,830	8,466,510	-	-	-	-	-	-
ANOS1	8,528,874	8,732,187	5,921,251	6,106,892	5,163,638	5,363,835	-	-
FAM9A	8,790,795	8,801,383	-	-	-	-	-	-
FAM9B	9,024,232	9,197,176	-	-	-	-	-	-
TBL1X	9,463,295	9,719,740	6,801,831	7,040,922	5,994,324	6,238,843	77,510,547	77,662,713
GPR143	9,725,413	9,765,965	7,027,703	7,095,421	6,242,296	6,276,654	152,781,921	152,808,651
LOC110257838	-	-	-	-	6,295,275	6,296,381	-	-
LOC109496425	-	-	7,099,539	7,124,318	-	-	-	-
SHROOM2	9,786,406	9,949,443	7,124,630	7,284,224	6,296,212	6,450,401	152,609,509	152,770,256
CLDN34	9,961,042	9,968,381	7,293,475	7,296,076	6,454,319	6,460,306	152,537,528	152,564,198
LOC101087815	-	-	7,315,456	7,316,115	-	-	-	-
WWC3	10,015,755	10,144,478	7,355,992	7,475,054	6,815,648	6,930,939	-	-
CLCN4	10,156,945	10,237,660	7,496,466	7,559,678	6,939,637	7,019,745	-	-
Mouse Ampliconic	-	-	-	-	-	-	***170,737,096***	***170,881,298***
MID1	10,445,310	11,111,264	7,719,203	8,316,004	7,235,386	7,906,049	169,685,199	169,990,798
HCCS	11,111,286	11,123,086	8,316,195	8,326,732	7,882,609	7,894,489	169,311,530	169,320,371
ARHGAP6	11,137,543	11,665,701	8,336,730	8,785,167	7,910,851	8,441,579	168,795,094	169,304,440
AMELX	11,293,413	11,300,761	8,439,198	8,453,188	8,055,246	8,063,457	169,175,044	169,187,284
MSL3	11,758,159	11,775,753	8,883,583	8,898,214	8,519,366	8,535,249	168,651,267	168,674,335
FRMPD4	12,137,329	12,724,523	8,941,173	9,799,966	8,625,039	9,451,433	167,471,301	168,577,241
LOC110257839	-	-	-	-	9,501,052	9,505,154	-	-

Table B3.3. Continued

Gene Name	Human Start	Human End	Cat Start	Cat End	Pig Start	Pig End	Mouse Start	Mouse End
Gm6744	-	-	-	-	-	-	167,445,442	167,446,609
PRPS2	12,791,355	12,824,227	9,844,030	9,880,152	9,505,034	9,531,044	167,346,320	167,383,106
TLR7	12,867,083	12,890,361	9,893,072	9,918,155	9,573,178	9,576,987	167,304,926	167,330,574
TLR8	12,906,620	12,923,169	9,935,203	9,951,074	9,551,284	9,616,762	167,241,123	167,264,329
TMSB4X	12,975,107	12,977,227	9,997,226	9,999,368	9,649,516	9,651,662	167,207,093	167,213,479
FAM9C	13,035,617	13,044,798	-	-	-	-	-	-
LOC105373133	13,044,544	13,086,975	-	-	-	-	-	-
ATXN3L	13,318,647	13,320,399	-	-	-	-	-	-
Gm8817	-	-	-	-	-	-	167,063,449	167,070,370
EGFL6	13,569,575	13,633,575	10,554,046	10,610,332	10,218,108	10,284,152	166,523,007	166,585,716
TCEANC	13,653,106	13,665,409	10,625,144	10,627,975	10,293,676	10,305,033	166,499,810	166,518,602
RAB9A	13,689,121	13,709,825	10,646,730	10,669,512	10,328,085	10,349,399	166,457,251	166,479,867
TRAPPC2	13,712,242	13,734,635	10,673,373	10,684,725	10,353,417	10,366,095	166,440,736	166,453,140
OFD1	13,734,713	13,769,361	10,684,924	10,726,453	10,361,403	10,415,502	166,390,033	166,440,704
GPM6B	13,770,922	13,938,824	10,748,223	10,897,814	10,429,496	10,582,768	166,238,943	166,389,033
GEMIN8	14,002,494	14,030,087	10,952,987	10,972,664	10,647,019	10,671,204	166,170,454	166,190,512
GLRA2	14,448,779	14,731,812	11,391,306	11,563,707	11,111,655	11,288,496	165,129,017	165,327,867
FANCB	14,690,863	14,873,255	11,668,259	11,693,030	11,386,360	11,408,958	164,980,524	164,997,272
MOSPD2	14,873,405	14,922,166	11,693,313	11,746,866	11,409,157	11,473,969	164,936,169	164,980,375
ASB9	15,243,987	15,270,476	12,064,094	12,091,544	11,800,546	11,828,251	164,497,903	164,539,752
ASB11	15,281,697	15,315,624	12,100,327	12,125,241	11,842,855	11,871,040	164,436,951	164,459,170
PIGA	15,319,451	15,335,554	12,127,895	12,140,825	11,871,746	11,886,519	164,419,685	164,433,916
FIGF	15,345,591	15,384,413	12,146,973	12,182,085	11,893,254	11,933,923	164,373,000	164,402,650
PIR	15,384,799	15,493,588	12,182,205	12,271,870	11,934,157	12,034,629	164,269,431	164,373,013
ACE2	15,494,525	15,602,069	12,330,854	12,374,725	12,099,849	12,152,204	164,139,342	164,188,418

Table B3.3. Continued

Gene Name	Human Start	Human End	Cat Start	Cat End	Pig Start	Pig End	Mouse Start	Mouse End
CXH4orf3	-	-	-	-	12,139,795	12,141,315	-	-
BMX	15,500,777	15,556,529	12,275,104	12,328,694	12,044,110	12,096,042	164,192,842	164,258,193
TMEM27	15,627,316	15,675,624	12,398,418	12,433,031	12,183,793	12,215,474	164,090,187	164,118,859
LOC101085781 (CA5BL)	-	-	12,440,795	12,466,881	***12,224,022***	***12,247,054***	-	-
Siah1b	-	-	-	-	-	-	164,070,703	164,076,493
CA5B	15,738,281	15,787,625	12,514,663	12,563,856	***12,295,393***	***12,332,283***	163,976,822	164,028,010
ZRSR2	15,789,723	15,825,718	12,569,212	12,597,132	12,336,635	12,365,071	163,935,443	163,958,702
AP1S2	15,825,806	15,855,014	12,599,333	12,627,781	12,367,785	12,397,672	163,908,706	163,933,666
Rnf138rt1	-	-	-	-	-	-	163,760,139	163,761,332
GRPR	16,123,301	16,153,518	12,883,373	12,917,047	12,654,059	12,711,417	163,513,904	163,549,736
MAGEB17	16,167,481	16,172,394	12,928,954	12,932,714	-	-	-	-
LOC101086788	-	-	13,225,739	13,226,878	-	-	-	-
LOC105261399	-	-	13,277,526	13,280,230	-	-	-	-
CTPS2	16,587,999	16,712,979	13,293,667	13,398,573	13,050,740	13,181,131	162,901,243	163,034,541
S100G	16,649,787	16,654,674	13,338,586	13,344,817	13,107,611	13,114,213	162,961,992	162,964,599
SYAP1	16,719,584	16,762,684	13,398,734	13,429,730	13,181,948	13,206,948	162,856,843	162,888,462
TXLNG	16,786,432	16,844,519	13,453,376	13,507,683	13,229,680	13,292,371	162,778,917	162,829,454
RBBP7	16,844,650	16,870,411	13,508,890	13,533,063	13,293,278	13,317,074	162,760,372	162,779,090
REPS2	16,946,691	17,194,274	13,611,280	13,833,465	13,413,362	13,658,907	162,411,952	162,643,705
NHS	17,375,420	17,735,994	14,004,032	14,347,483	13,854,950	14,209,085	161,833,290	162,159,441
SCML1	17,737,449	17,754,988	***14,349,325***	***14,366,188***	14,210,164	14,222,053	161,815,065	161,821,376
RAI2	17,800,049	17,861,346	14,393,607	14,453,640	14,246,776	14,633,071	161,717,036	161,779,494
BEND2	18,162,931	18,220,904	14,771,960	14,815,893	14,635,795	14,703,934	161,357,881	161,364,590
LOC108168453	-	-	-	-	-	-	161,297,856	161,334,722
SCML2	18,239,313	18,355,089	14,859,002	14,960,727	14,715,357	14,823,016	161,082,629	161,258,218

Table B3.3. Continued

Gene Name	Human Start	Human End	Cat Start	Cat End	Pig Start	Pig End	Mouse Start	Mouse End
CDKL5	18,425,605	18,653,629	15,029,852	15,243,399	14,858,826	15,096,969	160,784,308	160,994,681
LOC102899587	-	-	15,116,750	15,122,803	-	-	-	-
RS1	18,639,688	18,672,103	15,229,578	15,255,709	15,080,972	15,099,426	160,768,013	160,799,663
PPEF1	18,675,909	18,827,921	15,271,558	15,381,135	15,111,230	15,251,415	160,622,419	160,751,290
PHKA2	18,892,298	18,984,362	15,415,934	15,486,429	15,279,398	15,365,731	160,502,155	160,598,879
ADGRG2	18,989,307	19,122,956	15,492,264	15,614,222	15,370,327	15,518,746	160,390,649	160,498,076
LOC101089581	-	-	15,755,581	15,774,652	-	-	-	-
PDHA1	19,343,893	19,361,707	15,796,823	15,813,203	15,699,924	15,718,908	160,122,219	160,138,430
MAP3K15	19,360,058	19,515,459	15,811,748	15,936,106	15,716,898	15,867,377	159,988,433	160,123,351
SH3KBP1	19,533,965	19,887,626	15,952,149	16,288,071	15,887,131	16,229,212	159,627,028	159,975,920
BCLAF3	19,912,860	19,991,040	16,315,457	16,372,528	16,252,407	16,312,426	159,532,640	159,593,081
MAP7D2	20,006,713	20,116,996	16,385,043	16,491,221	16,327,718	16,423,546	159,414,524	159,498,958
EIF1AX	20,124,518	20,141,848	16,498,450	16,521,170	16,429,625	16,449,581	159,372,195	159,385,699
RPS6KA3	20,149,911	20,267,514	16,527,215	16,648,281	16,452,401	16,576,854	159,255,696	159,368,244
CNKSR2	21,374,335	21,654,695	17,816,263	18,115,397	17,561,114	17,838,034	157,821,573	158,043,353
KLHL34	21,655,349	21,658,387	18,115,305	18,120,112	17,838,070	17,841,718	157,818,435	157,821,060
SMPX	21,705,972	21,758,160	18,213,267	18,231,904	17,889,030	17,938,206	157,698,973	157,752,591
MBTPS2	21,839,538	21,885,423	18,321,450	18,362,813	18,013,360	18,054,250	157,547,822	157,598,715
YY2	21,855,987	21,858,727	18,337,990	18,341,436	-	-	157,566,119	157,568,985
SMS	21,940,573	21,994,837	18,423,545	18,477,918	18,118,332	18,169,396	157,443,852	157,492,046
PHEX	22,032,327	22,251,310	18,503,662	18,725,404	18,199,396	18,420,985	157,162,075	157,415,286
ZNF645	22,272,913	22,274,461	-	-	-	-	-	-
DDX53	22,999,961	23,002,089	19,497,659	19,510,658	19,164,076	19,169,280	-	-
PTCHD1	23,334,370	23,404,374	19,830,356	19,883,291	19,508,501	19,563,298	155,569,736	155,624,178
LOC101097703	-	-	20,032,463	20,032,949	-	-	-	-

Table B3.3. Continued

Gene Name	Human Start	Human End	Cat Start	Cat End	Pig Start	Pig End	Mouse Start	Mouse End
LOC101092829	-	-	20,037,757	20,038,541	-	-	-	-
PRDX4	23,664,260	23,686,399	20,117,734	20,140,798	19,800,153	19,819,575	155,323,918	155,340,749
ACOT9	23,701,055	23,743,290	20,147,901	20,174,189	19,835,674	19,864,570	155,262,443	155,297,654
SAT1	23,783,158	23,786,223	20,204,990	20,208,126	19,908,037	19,910,553	155,213,126	155,216,449
Cldn34b2	-	-	-	-	-	-	155,124,917	155,147,849
Cldn34b1	-	-	-	-	-	-	154,886,803	154,913,643
Gm15142	-	-	-	-	-	-	154,635,586	154,639,130
Samt2	-	-	-	-	-	-	154,575,228	154,579,330
Samt4	-	-	-	-	-	-	154,482,002	154,484,682
Gm5646	-	-	-	-	-	-	154,398,215	154,402,162
4930524N10Rik	-	-	-	-	-	-	154,339,225	154,343,392
Gm5645	-	-	-	-	-	-	154,130,867	154,134,052
Gm15140	-	-	-	-	-	-	154,109,633	154,120,685
Gm5947	-	-	-	-	-	-	154,079,612	154,082,604
Gm15143	-	-	-	-	-	-	***154,065,195***	***154,069,350***
Gm10057	-	-	-	-	-	-	***154,043,986***	***154,055,041***
4921511M17Rik	-	-	-	-	-	-	***154,022,780***	***154,033,827***
Samt1	-	-	-	-	-	-	154,001,590	154,012,640
Gm8787	-	-	-	-	-	-	79,330,513	79,358,068
Tsga8	-	-	-	-	-	-	83,486,674	83,487,924
Samt3	-	-	-	-	-	-	86,044,199	86,047,313
4930415L06Rik (PPP4RBL3like)	-	-	-	-	-	-	89,930,097	89,932,852
Gm5072	-	-	-	-	-	-	91,392,500	91,393,534
Gm14781	-	-	-	-	-	-	***91,624,249***	***91,635,671***
LOC108168446 (MAGEB2L)	-	-	-	-	-	-	***91,625,353***	***91,626,477***

Table B3.3. Continued

Gene Name	Human Start	Human End	Cat Start	Cat End	Pig Start	Pig End	Mouse Start	Mouse End
Mageb1	-	-	-	-	-	-	92,015,191	92,016,333
Gm5941	-	-	-	-	-	-	92,489,949	92,490,681
AU015836	-	-	-	-	-	-	93,968,656	93,975,470
Mouse Ampliconic	-	-	-	-	-	-	***94,706,195***	***94,816,711***
APOO	23,833,348	23,907,940	20,247,999	20,304,786	19,953,147	20,017,784	94,367,011	94,417,093
CXorf58	23,907,932	23,939,520	20,304,897	20,329,696	20,020,746	20,051,819	94,339,903	94,366,937
KLHL15	23,983,716	24,027,186	20,375,326	20,410,641	20,088,208	20,130,452	94,234,512	94,277,724
EIF2S3	24,054,948	24,078,810	20,423,874	20,442,946	20,129,708	20,162,881	94,188,709	94,212,651
ZFX	24,148,994	24,216,255	20,504,073	20,557,450	20,220,812	20,273,689	94,074,631	94,150,645
SUPT20HL2	24,310,862	24,313,315	-	-	-	-	-	-
SUPT20HL1	24,362,761	24,365,424	-	-	-	-	-	-
LOC100156694	-	-	-	-	20,366,396	20,433,395	-	-
PDK3	24,465,227	24,550,466	20,750,649	20,816,368	20,486,205	20,598,492	93,764,616	93,832,095
PCYT1B	24,558,087	24,672,862	20,833,107	20,968,217	20,605,769	20,743,081	93,654,863	93,749,951
LOC102167889 (RNF168)	-	-	-	-	20,654,157	20,659,555	-	-
POLA1	24,693,833	24,996,986	20,981,713	21,286,059	20,756,674	21,064,265	93,304,766	93,632,219
ARX	25,003,694	25,015,948	21,292,456	21,310,887	21,067,631	21,079,850	93,286,507	93,298,357
MAGEB18	26,138,343	26,140,736	22,115,373	22,265,770	22,082,917	22,084,040	92,118,879	92,599,572
MAGEB6	26,192,440	26,195,646	-	-	22,126,833	22,128,116	***91,331,905***	***91,332,897***
LOC100513448 (MAGEB4L)	-	-	-	-	22,170,175	22,171,346	-	-
FELCATV1R3	-	-	21,603,090	21,603,992	-	-	-	-
LOC105261400 (MAGEB4L)	-	-	22,294,290	22,295,724	-	-	-	-
LOC111558670 (MAGEB17L)	-	-	22,330,125	22,331,507	-	-	-	-
MAGEB5	26,216,169	26,218,270	22,344,146	22,344,812	22,172,038	22,176,128	***91,779,608***	***91,782,976***
LOC493792 (MAGEB5L)	-	-	22,348,237	22,356,721	22,188,608	22,189,561	90,878,355	90,893,134

Table B3.3. Continued

Gene Name	Human Start	Human End	Cat Start	Cat End	Pig Start	Pig End	Mouse Start	Mouse End
PPP4R3C	27,460,211	27,463,341	23,399,287	23,402,011	23,301,323	23,303,743	89,752,268	89,755,491
LOC100157863 (APOPT1)	-	-	-	-	22,381,742	22,382,524	-	-
LOC100514934 (RPL39L)	-	-	-	-	23,259,680	23,259,840	-	-
DCAF8L2	27,746,809	27,748,821	-	-	23,566,754	23,569,786	89,403,848	89,409,689
MAGEB10	27,807,990	27,823,014	23,699,598	23,703,650	23,618,483	23,621,851	-	-
MAGEB10L	-	-	23,711,760	23,713,477	23,629,744	23,632,620	-	-
LOC100156962 (MAGEB10L)	-	-	-	-	23,652,358	23,653,374	-	-
LOC102159844 (MAGEB4L)	-	-	-	-	25,895,424	26,030,203	-	-
LOC493964 (MAGE)	-	-	23,749,406	23,751,663	-	-	-	-
DCAF8L1	27,977,993	27,981,449	-	-	-	-	-	-
IL1RAPL1	28,587,564	29,956,350	24,698,982	25,853,110	24,408,092	25,811,129	86,747,242	88,115,623
MAGEB2 (duplicate of MAGEB4)	30,215,558	30,220,089	-	-	-	-	-	-
LOC109496467 (MAGEB2L)	-	-	26,047,369	26,054,530	-	-	-	-
MAGEB3	30,230,436	30,237,493	26,047,369	26,070,090	26,044,502	26,055,412	-	-
MAGEB4	30,241,940	30,244,193	26,083,553	26,088,438	26,044,526	26,076,729	86,250,256	86,256,219
MAGEB1(duplicate of MAGEB4)	30,243,731	30,252,038	-	-	-	-	-	-
NR0B1	30,304,422	30,309,399	26,125,426	26,131,512	26,117,874	26,122,951	86,191,775	86,195,947
LOC109496468	-	-	26,204,512	26,480,297	-	-	-	-
CXorf21	30,558,824	30,577,916	26,390,148	26,406,290	26,363,808	26,384,933	85,776,968	85,891,499
GK	30,653,348	30,731,462	26,481,219	26,561,588	26,470,422	26,558,892	85,701,937	85,776,819
TAB3	30,827,442	30,889,394	26,650,274	26,739,846	26,632,547	26,729,039	85,573,851	85,634,469
FTHL17	31,071,241	31,072,053	26,943,996	26,944,900	26,955,098	26,956,111	85,249,677	85,270,291
DMD	31,119,219	33,339,609	26,988,212	29,005,177	27,028,223	29,650,728	82,814,664	85,205,050
FAM47A	34,129,752	34,132,330	-	-	-	-	-	-
E3-like	-	-	30,180,347	30,187,170	30,364,427	30,367,647	81,419,575	81,458,122

Table B3.3. Continued

Gene Name	Human Start	Human End	Cat Start	Cat End	Pig Start	Pig End	Mouse Start	Mouse End
TMEM47	34,627,064	34,657,288	30,396,753	30,426,634	30,584,044	30,615,233	81,070,644	81,097,875
FAM47B	34,942,796	34,944,917	-	-	-	-	-	-
LOC100623332 (MAGEB16L)	-	-	-	-	31,357,792	31,721,181	-	-
LOC100522230 (MAGEB16L)	-	-	-	-	31,573,305	31,574,593	-	-
LOC100522411 (MAGEB16L)	-	-	-	-	31,592,308	31,597,919	-	-
LOC100623432 (MAGEB16L)	-	-	-	-	***31,752,126***	***31,753,755***	-	-
LOC100520818 (MAGEB16L)	-	-	-	-	***31,759,602***	***31,780,300***	-	-
MAGEB16	35,798,317	35,803,753	31,497,919	31,501,538	30,874,573	30,945,913	79,623,231	79,671,435
CFAP47	35,919,734	36,385,319	31,629,513	32,213,972	31,879,143	32,331,543	79,482,568	79,517,285
LOC105373154 (MIB-like)	36,902,398	36,959,104	-	-	-	-	-	-
H2A-Bbd type 1-like	-	-	32,523,326	32,523,883	32,692,523	32,693,072	-	-
FAM47C	37,008,359	37,011,666	32,779,211	32,780,853	33,004,811	33,006,091	78,737,763	78,739,410
PRRG1	37,349,275	37,457,295	32,686,315	32,996,245	33,064,499	33,232,556	78,449,610	78,583,912
LOC110257841 (H2A-Bbd type 1-like)	-	-	-	-	33,299,003	33,299,753	-	-
LOC100624295 (H2A-Bbd type 1-like)	-	-	-	-	33,307,648	33,308,398	-	-
LANCL3	37,571,569	37,684,463	33,177,349	33,273,100	33,426,731	33,543,506	9,199,973	9,268,085
XK	37,685,766	37,732,130	33,282,834	33,351,222	33,544,287	33,602,576	9,272,784	9,313,245
LOC111558783	-	-	33,312,823	33,313,387	-	-	-	-
H2al1m	-	-	-	-	-	-	9,283,764	9,284,288
H2al1k	-	-	-	-	-	-	9,350,599	9,351,137
CYBB	37,780,017	37,813,461	33,385,703	33,419,837	33,654,656	33,678,346	9,435,252	9,487,766
H2al1o	-	-	-	-	-	-	9,571,960	9,572,391
DYNLT3	37,838,836	37,847,636	33,444,706	33,452,472	33,711,993	33,724,596	9,654,270	9,662,983
SYTL5	37,888,915	38,128,820	33,506,964	33,859,182	33,798,296	34,066,127	9,747,559	9,998,864
HYPM	37,990,817	37,991,317	33,634,621	33,635,188	33,889,233	33,890,290	9,846,947	9,847,500

Table B3.3. Continued

Gene Name	Human Start	Human End	Cat Start	Cat End	Pig Start	Pig End	Mouse Start	Mouse End
LOC100736633 (H2A-Bbd type 1-like)	-	-	-	-	33,894,983	33,895,366	-	-
H2al3	-	-	-	-	-	-	9,849,703	9,850,252
SRPX	38,149,335	38,220,924	33,878,250	33,990,363	34,075,577	34,198,462	10,037,977	10,117,661
RPGR	38,269,163	38,327,564	34,000,088	34,082,719	34,238,961	34,522,814	10,158,216	10,216,795
OTC	38,352,483	38,421,450	34,105,699	34,167,849	34,322,398	34,391,711	10,252,296	10,321,024
LOC101083747 (FTL-like)	-	-	34,224,027	34,292,046	-	-	-	-
TSPAN7	38,561,478	38,688,918	34,292,066	34,417,782	34,522,899	34,660,216	10,485,116	10,596,604
MID1IP1	38,801,432	38,806,532	34,498,781	34,501,121	34,763,360	34,766,160	10,717,365	10,719,702
H2al1a	-	-	-	-	-	-	***11,299,257***	***11,299,757***
H2a1b	-	-	-	-	-	-	***11,302,432***	***11,302,921***
H2a1c	-	-	-	-	-	-	***11,305,655***	***11,305,972***
H2a1d	-	-	-	-	-	-	***11,308,824***	***11,309,141***
H2al1e	-	-	-	-	-	-	***11,311,934***	***11,312,427***
H2a1f	-	-	-	-	-	-	***11,315,158***	***11,315,475***
H2a1g	-	-	-	-	-	-	***11,318,256***	***11,318,764***
H2a1h	-	-	-	-	-	-	***11,321,453***	***11,321,894***
H2a1i	-	-	-	-	-	-	***11,324,659***	***11,324,976***
H2a1j	-	-	-	-	-	-	***11,327,822***	***11,328,151***
BCOR	40,051,246	40,177,390	35,544,903	35,658,110	35,768,013	35,891,013	12,036,737	12,160,528
LOC107985687 (EIF2L)	40,101,736	40,106,486	-	-	-	-	-	-
LOC106506945 (SRRM1L)	-	-	-	-	35,880,111	35,904,287	-	-
LOC110257671	-	-	-	-	36,042,052	36,140,080	-	-
ATP6AP2	40,580,964	40,606,637	35,989,837	36,013,165	36,271,154	36,300,240	12,587,759	12,617,051
MPC1L	40,623,566	40,624,139	36,038,180	36,038,906	36,325,447	36,331,779	-	-
CXorf38	40,626,921	40,647,678	36,044,140	36,061,506	36,330,684	36,351,903	12,654,879	12,675,377

Table B3.3. Continued

Gene Name	Human Start	Human End	Cat Start	Cat End	Pig Start	Pig End	Mouse Start	Mouse End
MED14	40,649,543	40,736,122	36,064,001	36,131,139	36,354,057	36,424,837	12,675,371	12,762,106
MED14OS	40,735,396	40,738,701	36,131,023	36,313,801	-	-	-	-
USP9X	41,085,420	41,236,579	36,459,686	36,591,312	36,736,049	36,888,702	***13,070,995***	***13,173,328***
DDX3X	41,333,308	41,364,472	36,683,612	36,700,297	36,988,031	37,005,496	13,280,496	13,293,988
NYX	41,447,460	41,476,414	36,802,939	36,806,275	37,080,197	37,105,513	13,461,095	13,489,321
CASK	41,514,934	41,923,525	36,864,960	37,222,095	37,166,973	37,535,723	13,517,080	13,846,784
LOC110257750	-	-	-	-	37,168,973	37,172,889	-	-
GPR34	41,686,543	41,697,277	37,010,409	37,019,112	37,306,490	37,313,626	13,623,413	13,640,859
GPR82	41,724,155	41,730,135	37,034,019	37,041,811	37,328,672	37,334,144	13,658,153	13,667,433
PPP1R2C	42,777,368	42,778,247	37,930,416	37,931,231	38,292,829	38,293,616	-	-
H2al1n	-	-	-	-	-	-	14,211,148	14,211,657
Cypt1	-	-	-	-	-	-	16,522,869	16,523,691
MAOA	43,654,907	43,746,824	38,669,875	38,744,690	38,930,452	39,006,220	16,619,698	16,687,812
MAOB	43,766,610	43,882,475	38,762,524	38,874,356	39,025,973	39,144,389	16,709,281	16,817,366
NDP	43,948,776	43,973,675	38,946,574	38,998,125	39,208,115	39,235,879	16,885,521	16,911,774
EFHC2	44,146,254	44,343,677	39,144,009	39,345,247	39,407,390	39,597,537	17,132,049	17,320,470
FUNDC1	44,523,639	44,542,975	39,503,089	39,515,318	39,726,194	39,737,663	17,519,967	17,572,325
DUSP21	44,844,003	44,844,888	39,721,372	39,722,677	39,929,707	39,930,539	18,145,870	18,146,694
KDM6A	44,873,175	45,112,612	39,746,470	39,955,494	39,953,764	40,154,197	18,162,364	18,279,936
CXorf36	45,148,373	45,200,901	39,995,149	40,041,707	40,179,289	40,223,011	18,414,167	18,461,421
KRBOX4	46,447,292	46,473,669	-	-	-	-	-	-
ZNF674	46,497,725	46,545,459	41,009,784	41,045,467	41,090,146	41,118,627	-	-
Tex13c3	-	-	-	-	-	-	19,479,424	19,481,319
CHST7	46,573,687	46,598,496	41,091,066	41,114,709	41,192,993	41,234,654	20,058,962	20,097,524
SLC9A7	46,599,251	46,759,172	41,115,619	41,249,907	41,233,308	41,399,032	20,098,851	20,291,968

Table B3.3. Continued

Gene Name	Human Start	Human End	Cat Start	Cat End	Pig Start	Pig End	Mouse Start	Mouse End
RP2	46,836,912	46,882,358	41,327,544	41,375,360	41,488,866	41,534,810	20,364,481	20,405,653
JADE3	46,912,276	47,061,242	41,407,097	41,543,947	41,579,565	41,714,705	20,425,327	20,519,939
LOC110257673 (PRB2-like)	-	-	-	-	41,579,893	41,581,086	-	-
RGN	47,078,299	47,093,314	41,566,278	41,579,275	41,725,660	41,749,736	20,549,766	20,562,089
NDUFB11	47,142,216	47,145,210	41,603,687	41,606,341	41,770,929	41,773,879	20,615,326	20,617,564
RBM10	47,145,196	47,186,815	41,606,734	41,637,288	41,773,232	41,803,656	20,617,503	20,650,905
UBA1	47,190,800	47,215,128	41,644,448	41,668,327	41,810,726	41,832,818	20,658,302	20,683,179
CDK16	47,217,881	47,229,997	41,671,670	41,682,706	41,836,268	41,848,580	20,687,954	20,699,880
USP11	47,232,915	47,248,328	41,686,637	41,702,578	41,851,534	41,866,357	20,702,326	20,720,539
ZNF157	47,370,600	47,413,699	41,810,153	41,851,923	41,905,319	41,921,484	-	-
ZNF41	47,444,691	47,483,671	41,865,169	41,911,822	41,931,036	41,984,439	-	-
Gm39500	-	-	-	-	-	-	20,775,461	20,783,201
ARAF	47,561,100	47,571,921	41,987,525	41,999,008	42,078,202	42,090,250	20,797,784	20,860,521
SYN1	47,571,901	47,619,857	41,999,004	42,052,055	42,090,245	42,142,254	20,860,511	20,920,918
TIMP1	47,582,291	47,586,791	42,009,536	42,013,541	42,103,251	42,106,289	20,870,166	20,874,737
CFP	47,624,213	47,630,305	42,056,379	42,061,782	42,145,754	42,152,879	20,925,454	20,931,555
ELK1	47,635,520	47,650,604	42,066,948	42,081,493	42,157,980	42,174,964	20,933,395	20,950,608
UXT	47,651,792	47,659,180	42,082,681	42,091,622	42,176,142	42,184,131	20,951,665	20,962,034
CXXC1P1	-	-	42,125,508	42,142,686	42,227,638	42,232,777	-	-
ZNF81	47,836,864	47,925,626	42,198,548	42,288,916	42,246,548	42,372,191	-	-
ZNF182	47,974,851	48,003,995	42,377,460	42,408,126	42,444,227	42,473,308	21,026,184	21,062,083
SPACA5	48,004,336	48,009,736	42,412,167	42,417,462	42,476,719	42,478,783	21,068,488	21,077,959
Zfp300	-	-	-	-	-	-	21,079,150	21,089,700
Ssxa1 (Ssxb paralog)	-	-	-	-	-	-	21,118,414	21,122,088
Gm5932	-	-	-	-	-	-	21,169,950	21,179,036

Table B3.3. Continued

Gene Name	Human Start	Human End	Cat Start	Cat End	Pig Start	Pig End	Mouse Start	Mouse End
4930453H23Rik	-	-	-	-	-	-	21,206,580	21,221,672
ZNF630	48,058,178	48,071,658	42,479,945	42,491,478	42,518,203	42,540,677	-	-
SPACA5B (SPACA5 dup)	48,130,657	48,132,613	-	-	-	-	-	-
SSX5	48,177,982	48,196,798	-	-	-	-	8,803,691	8,811,422
SSX1	48,255,317	48,267,444	42,522,876	42,531,522	42,559,714	42,567,979	-	-
SSX3	48,346,428	48,356,753	-	-	-	-	-	-
SSX4	***48,383,527***	***48,393,343***	-	-	-	-	-	-
SSX4B	***48,402,082***	***48,411,910***	-	-	-	-	-	-
4930402K13Rik (Fam47c duplicate)	-	-	-	-	-	-	9,104,562	9,106,342
Fthl17f	-	-	-	-	-	-	9,063,090	9,063,822
Fthl17e	-	-	-	-	-	-	9,033,486	9,034,333
Fthl17d	-	-	-	-	-	-	8,985,940	8,986,672
Fthl17c	-	-	-	-	-	-	8,975,718	8,976,559
Fthl17b	-	-	-	-	-	-	8,962,134	8,962,975
Gm5751	-	-	-	-	-	-	8,871,568	8,875,407
Gm6592	-	-	-	-	-	-	8,843,994	8,849,607
Ssx9	-	-	-	-	-	-	8,743,792	8,756,633
Ssxb8	-	-	-	-	-	-	8,684,843	8,690,891
Ssxb7	-	-	-	-	-	-	8,639,607	8,645,722
Ssxb3	-	-	-	-	-	-	8,583,404	8,591,810
Ssxb6	-	-	-	-	-	-	8,542,604	8,548,255
Gm14459	-	-	-	-	-	-	8,513,543	8,526,726
Mouse Ampliconic	-	-	-	-	-	-	***8,480,304***	***8,499,226***
Ssxb2	-	-	-	-	-	-	8,454,343	8,461,726
Ssxb1	-	-	-	-	-	-	8,413,305	8,422,158

Table B3.3. Continued

Gene Name	Human Start	Human End	Cat Start	Cat End	Pig Start	Pig End	Mouse Start	Mouse End
Ssxb9	-	-	-	-	-	-	8,366,978	8,375,388
Ssxb10	-	-	-	-	-	-	8,327,424	8,336,235
SLC38A5	48,458,537	48,470,256	42,545,288	42,558,731	42,590,608	42,601,537	8,271,381	8,280,179
FTSJ1	48,476,021	48,486,364	42,560,905	42,570,749	42,609,560	42,620,198	8,238,668	8,252,406
PORCN	48,508,954	48,520,814	42,611,838	42,623,667	42,653,696	42,669,526	8,193,846	8,206,546
EBP	48,521,776	48,528,716	42,624,082	42,630,249	42,670,592	42,678,571	8,185,329	8,193,512
TBC1D25	48,539,687	48,562,609	42,631,742	42,647,931	42,682,267	42,696,579	8,154,470	8,176,181
RBM3	48,574,353	48,581,165	42,653,365	42,656,949	42,704,019	42,707,587	8,138,975	8,145,880
WDR13	48,597,492	48,605,194	42,675,596	42,683,048	42,715,097	42,730,200	8,123,301	8,132,858
WAS	48,683,753	48,691,427	42,760,775	42,768,590	42,774,597	42,793,055	8,081,466	8,090,491
SUV39H1	48,695,554	48,709,016	42,771,556	42,786,344	42,797,309	42,811,457	8,061,171	8,074,760
GLOD5	48,761,572	48,773,648	42,848,010	42,855,293	42,877,034	42,889,285	8,004,201	8,018,492
GATA1	48,786,540	48,794,311	42,867,799	42,874,829	42,898,431	42,905,939	7,959,260	7,976,682
HDAC6	48,801,377	48,824,982	42,880,987	42,906,239	42,917,911	42,941,894	7,930,116	7,948,016
ERAS	48,826,513	48,829,869	42,909,793	42,911,055	42,942,033	42,946,715	7,924,276	7,928,607
PCSK1N	48,831,092	48,835,638	42,911,958	42,917,448	42,947,688	42,952,222	7,919,822	7,924,410
TIMM17B	48,893,447	48,898,143	42,933,353	42,937,503	42,970,586	42,976,287	7,899,336	7,907,652
PQBP1	48,897,862	48,903,145	42,937,601	42,941,626	42,976,086	42,981,732	7,894,519	7,899,269
SLC35A2	48,903,180	48,911,958	42,941,684	42,950,709	42,981,801	42,992,091	7,884,034	7,894,492
PIM2	48,913,182	48,919,136	42,951,920	42,957,023	42,993,547	42,998,575	7,878,306	7,883,432
OTUD5	48,922,024	48,958,386	42,959,969	42,988,125	43,001,491	43,029,471	7,841,346	7,876,626
KCND1	48,960,983	48,972,099	42,991,245	42,999,501	43,032,355	43,043,310	7,823,759	7,838,278
GRIPAP1	48,973,720	49,002,264	43,001,341	43,023,786	43,043,613	43,070,363	7,789,961	7,820,567
TFE3	49,028,726	49,043,517	43,057,177	43,068,511	43,100,375	43,116,377	7,762,537	7,775,202
CCDC120	49,053,432	49,069,858	43,079,171	43,093,927	43,128,938	43,146,777	7,731,714	7,741,324

Table B3.3. Continued

Gene Name	Human Start	Human End	Cat Start	Cat End	Pig Start	Pig End	Mouse Start	Mouse End
PRAF2	49,071,156	49,074,045	43,094,355	43,097,227	43,147,473	43,150,350	7,728,571	7,731,063
WDR45	49,074,433	49,101,121	43,097,613	43,102,974	43,150,691	43,155,981	7,721,983	7,728,201
GPKOW	49,113,385	49,124,547	43,134,333	43,148,937	43,177,264	43,192,021	7,697,125	7,710,265
LOC101100947	-	-	43,153,840	43,155,056	-	-	-	-
MAGIX	49,162,569	49,168,157	43,160,994	43,164,969	43,206,680	43,212,889	7,673,166	7,681,251
PLP2	49,171,837	49,175,120	43,168,757	43,171,669	43,215,372	43,218,318	7,667,941	7,671,390
PRICKLE3	49,174,802	49,186,400	43,171,673	43,180,573	43,218,271	43,230,368	7,657,379	7,668,186
SYP	49,187,812	49,200,202	43,182,083	43,194,195	43,232,190	43,246,764	7,638,580	7,653,256
CACNA1F	49,205,063	49,233,404	43,197,672	43,222,747	43,250,449	43,284,458	7,607,070	7,635,196
CCDC22	49,235,467	49,250,526	43,224,062	43,237,373	43,285,232	43,303,865	7,593,809	7,605,420
FOXP3	49,250,436	49,266,505	43,237,305	43,256,645	43,303,777	43,328,164	7,579,676	7,595,243
PPP1R3F	49,269,822	49,301,465	43,256,161	43,272,133	43,328,199	43,344,503	7,558,562	7,574,281
GAGE10	49,303,646	49,319,844	43,299,794	43,308,070	-	-	-	-
GAGE12J	49,322,030	49,329,387	-	-	-	-	-	-
GAGE13	49,331,603	49,338,952	-	-	-	-	-	-
GAGE12B	49,341,183	49,529,921	-	-	-	-	-	-
GAGE12C	***49,532,177***	***49,539,541***	-	-	-	-	-	-
GAGE12D	***49,541,722***	***49,549,107***	-	-	-	-	-	-
GAGE12E	***49,551,289***	***49,558,649***	-	-	-	-	-	-
GAGE12F	***49,560,808***	***49,568,208***	-	-	-	-	-	-
GAGE12G	***49,570,400***	***49,577,757***	-	-	-	-	-	-
GAGE12H	***49,579,949***	***49,587,304***	-	-	-	-	-	-
GAGE2A	***49,589,515***	***49,596,827***	-	-	-	-	-	-
GAGE1	***49,598,992***	***49,608,536***	-	-	-	-	-	-
PAGE1	49,687,450	49,696,019	-	-	-	-	-	-

Table B3.3. Continued

Gene Name	Human Start	Human End	Cat Start	Cat End	Pig Start	Pig End	Mouse Start	Mouse End
PAGE4	49,829,260	49,834,264	43,324,947	43,475,715	43,489,559	43,496,654	-	-
USP27X	49,879,867	49,882,565	43,520,451	43,527,606	43,553,454	43,557,247	7,372,591	7,375,830
CLCN5	49,922,615	50,099,235	43,583,492	43,755,211	43,779,016	43,809,472	7,153,810	7,319,359
AKAP4	50,190,768	50,201,013	43,851,593	43,868,322	43,885,584	43,895,142	7,067,515	7,078,609
CCNB3	50,202,713	50,351,914	43,885,334	43,951,926	43,922,729	43,988,939	6,979,652	7,041,635
DGKK	50,365,407	50,470,850	43,988,268	44,151,488	44,001,172	44,180,142	6,779,306	6,948,363
SHROOM4	50,575,534	50,814,167	44,262,882	44,534,553	44,301,956	44,543,969	6,399,940	6,637,454
BMP15	50,910,735	50,916,641	44,645,465	44,650,172	44,613,122	44,620,575	6,314,106	6,320,723
LOC111558700 (Formin-like protein 4)	-	-	***44,771,130***	***44,776,751***	-	-	-	-
NUDT10	51,332,231	51,337,525	***44,776,919***	***44,782,971***	44,773,570	44,779,568	6,168,696	6,173,044
CXorf67	51,406,915	51,408,838	***44,897,331***	***44,902,505***	44,858,057	45,074,373	6,081,218	6,083,420
NUDT11	51,490,011	51,496,607	***44,936,298***	***44,942,817***	44,912,160	44,918,977	6,047,507	6,054,751
LOC109496513 (FMNL4L)	-	-	***44,942,817***	***44,947,984***	-	-	-	-
CENPVL3	51,619,054	51,711,363	***45,029,958***	***45,031,028***	-	-	-	-
CENPVL2	***51,681,212***	***51,682,831***	-	-	-	-	-	-
CENPVL1	***51,710,512***	***51,712,131***	-	-	-	-	-	-
LOC101088142	-	-	***45,159,674***	***45,160,706***	-	-	-	-
Btbd35f7	-	-	-	-	-	-	5,669,067	5,671,026
Mycs	-	-	-	-	-	-	5,466,904	5,469,265
Btbd35f27	-	-	-	-	-	-	***4,952,135***	***4,954,077***
Btbd35f4	-	-	-	-	-	-	***4,800,411***	***4,802,359***
Btbd35f17	-	-	-	-	-	-	***4,370,636***	***4,372,595***
Btbd35f20	-	-	-	-	-	-	***4,289,286***	***4,291,245***
Btbd35f28	-	-	-	-	-	-	***4,196,576***	***4,198,535***
Btbd35f3	-	-	-	-	-	-	***4,037,674***	***4,039,622***

Table B3.3. Continued

Gene Name	Human Start	Human End	Cat Start	Cat End	Pig Start	Pig End	Mouse Start	Mouse End
Btbd35f16	-	-	-	-	-	-	***3,955,043***	***3,957,002***
Btbd35f10	-	-	-	-	-	-	***3,750,931***	***3,752,885***
Btbd35f18	-	-	-	-	-	-	***3,700,233***	***3,702,192***
Btbd35f11	-	-	-	-	-	-	***3,441,731***	***3,443,690***
Btbd35f23	-	-	-	-	-	-	***3,076,920***	***3,078,823***
GSPT2	***51,743,385***	***51,746,232***	***45,252,788***	***45,255,337***	***45,193,336***	***45,195,939***	94,636,069	94,638,562
MAGED1	***51,802,998***	***51,902,354***	***45,392,408***	***45,399,213***	***45,258,761***	***45,329,282***	94,535,474	94,542,074
MAGED4B	***52,061,827***	***52,069,272***	-	-	-	-	-	-
MAGED4	***52,184,823***	***52,192,268***	***45,599,413***	***45,606,548***	-	-	-	-
XAGE2	***52,368,996***	***52,375,683***	-	-	-	-	-	-
XAGE1B	***52,495,668***	***52,500,815***	-	-	-	-	-	-
XAGE1E (XAGE1E)	***52,512,074***	***52,517,221***	-	-	-	-	-	-
SSX7	***52,644,061***	***52,654,900***	-	-	-	-	-	-
SSX2	***52,696,896***	***52,707,227***	-	-	-	-	-	-
SSX2B	***52,751,204***	***52,761,536***	-	-	-	-	-	-
SPANXN5	***52,796,146***	***52,797,348***	-	-	-	-	-	-
XAGE5	***52,811,295***	***52,818,301***	-	-	-	-	-	-
XAGE3	***52,862,528***	***52,868,140***	-	-	-	-	-	-
FAM156B	***52,891,599***	***52,908,560***	-	-	-	-	-	-
Gm15093 (Ott-like duplicate)	-	-	-	-	-	-	***149,272,377***	***149,305,003***
Gm15085 (Ott-like duplicate)	-	-	-	-	-	-	***149,447,617***	***149,487,784***
Gm10439	-	-	-	-	-	-	***149,594,601***	***149,636,621***
Gm15097 (Ott-like duplicate)	-	-	-	-	-	-	***149,784,424***	***149,826,694***
Gm15091 (Luzp4-like duplicate)	-	-	-	-	-	-	***149,941,387***	***149,984,982***
Gm20153	-	-	-	-	-	-	150,157,426	150,177,319

Table B3.3. Continued

Gene Name	Human Start	Human End	Cat Start	Cat End	Pig Start	Pig End	Mouse Start	Mouse End
Gm5946 (Bmi1-like)	-	-	-	-	-	-	150,220,899	150,223,351
Gm15104 (FAM156ABL)	-	-	-	-	-	-	150,311,924	150,336,719
FAM156A	52,947,254	52,995,472	Xrand	Xrand	***45,643,214***	***45,643,750***	150,397,773	150,467,331
LOC111558733	-	-	***45,708,172***	***45,712,838***	-	-	-	-
LOC102899730	-	-	***45,961,985***	***45,967,691***	-	-	-	-
Mouse Ampliconic	-	-	-	-	-	-	***152,409,657***	***152,438,240***
GPR173	53,049,324	53,080,615	46,084,470	46,108,219	45,883,738	45,908,197	152,343,598	152,368,966
TSPYL2	53,082,327	53,088,546	46,110,557	46,117,592	45,909,688	45,916,859	152,336,852	152,342,503
KDM5C	53,159,735	53,225,422	46,173,315	46,206,027	45,968,812	46,003,215	152,233,197	152,279,102
IQSEC2	53,225,784	53,321,350	46,214,037	46,290,383	46,011,386	46,094,978	152,144,183	152,225,237
SMC1A	53,374,149	53,422,728	46,330,306	46,377,254	46,143,738	46,202,187	152,016,428	152,061,973
Gm15266	-	-	-	-	-	-	152,110,432	152,111,413
RIBC1	53,422,782	53,431,120	46,377,407	46,392,446	46,202,213	46,218,691	152,004,580	152,016,995
HSD17B10	53,431,258	53,434,376	46,392,584	46,395,073	46,218,540	46,220,942	152,001,896	152,004,442
HUWE1	53,532,096	53,686,729	46,477,471	46,644,030	46,281,958	46,444,315	151,800,807	151,935,417
L1TD1	-	-	-	-	46,349,921	46,352,855	-	-
LOC106506967 (SMC1AL)	-	-	-	-	46,568,746	46,570,505	-	-
FCAB-31J21-174L20_Transcript1	-	-	***46,642,669****	***46,837,270****	-	-	-	-
Cat Ampliconic	-	-	***46,815,413***	***46,832,194***	-	-	-	-
PHF8	53,936,680	54,048,935	46,892,029	46,990,011	46,648,164	46,761,192	151,520,624	151,633,859
FAM120C	54,068,324	54,183,281	47,025,348	47,145,956	46,777,904	46,904,232	151,342,932	151,474,136
WNK3	54,192,823	54,358,695	47,153,760	47,342,844	46,923,587	47,096,265	151,168,778	151,320,194
TSR2	54,440,420	54,445,297	47,433,846	47,438,870	47,186,610	47,192,465	151,087,094	151,096,543
FGD1	54,445,454	54,496,166	47,439,018	47,481,215	47,191,838	47,231,656	151,046,118	151,089,686
GNL3L	54,530,211	54,567,287	47,508,287	47,534,367	47,256,467	47,284,520	150,983,133	151,017,322

Table B3.3. Continued

Gene Name	Human Start	Human End	Cat Start	Cat End	Pig Start	Pig End	Mouse Start	Mouse End
Pig Ampliconic	-	-	-	-	***47,435,127***	***47,474,791***	-	-
ITIH6	54,748,899	54,798,240	-	-	47,537,909	47,584,419	150,826,917	150,857,407
LOC106506966	-	-	-	-	47,552,766	47,561,116	-	-
MAGED2	54,807,599	54,816,015	47,800,318	47,808,578	-	-	150,806,421	150,814,340
TRO	54,920,563	54,931,436	47,930,092	47,940,454	47,759,334	47,772,014	150,644,710	150,657,583
PFKFB1	54,933,134	54,998,534	47,941,804	48,079,099	47,771,751	47,859,079	150,588,258	150,643,878
APEX2	55,000,323	55,007,873	48,080,375	48,090,338	47,862,682	47,870,579	150,503,629	150,588,208
ALAS2	55,009,055	55,031,064	48,091,113	48,117,734	47,871,519	47,896,041	150,547,417	150,570,622
PAGE2B	55,027,927	55,078,903	-	-	47,916,076	47,919,904	-	-
PAGE2	55,089,052	55,092,836	-	-	-	-	-	-
FAM104B	55,143,102	55,161,195	-	-	-	-	-	-
MTRNR2L10	55,181,391	55,182,511	-	-	-	-	-	-
PAGE5	55,220,347	55,224,108	-	-	-	-	-	-
PAGE3	55,257,833	55,264,916	-	-	-	-	-	-
Pig Ampliconic	-	-	-	-	***48,022,185***	***48,041,616***	-	-
MAGEH1	***55,452,089***	***55,453,568***	***48,439,547***	***48,440,970***	48,080,533	48,081,983	153,036,166	153,037,563
USP51	55,484,524	55,489,875	***48,383,283***	***48,387,342***	48,047,658	48,051,941	153,006,379	153,009,913
FOXR2	55,623,400	55,626,188	-	-	48,174,318	48,181,106	153,114,581	153,132,861
RRAGB	55,717,677	55,760,209	48,597,625	48,648,828	48,286,584	48,330,196	153,139,792	153,190,192
KLF8	55,908,123	56,291,531	48,778,131	49,024,463	48,663,063	48,752,359	153,225,339	153,397,886
UBQLN2	56,563,593	56,567,010	49,292,451	49,294,852	48,948,062	48,951,392	153,498,232	153,501,562
Cypt3	-	-	-	-	-	-	153,558,593	153,559,435
Kctd12b	-	-	-	-	-	-	153,685,154	153,696,386
NBDY	56,729,219	56,819,178	***49,419,354***	***49,534,625***	***49,156,877***	***49,230,907***	153,723,554	153,741,296
SPIN3	56,976,370	56,995,801	-	-	49,351,952	49,356,901	-	-

Table B3.3. Continued

Gene Name	Human Start	Human End	Cat Start	Cat End	Pig Start	Pig End	Mouse Start	Mouse End
FCAB-205P3...127P13_Transcript1			***49,467,101***	***49,468,645***				
SPIN2B	57,112,142	57,121,905	***49,821,100***	***49,844,334***	49,492,205	49,507,749	153,832,001	153,834,243
FAAH2	57,121,861	57,489,196	-	-	49,685,285	49,744,246	-	-
SPIN2A	57,134,259	57,138,688	-	-	-	-	-	-
LOC100511853 (GMCL1L1)	-	-	-	-	49,726,706	49,729,782	-	-
ZXDB	57,591,836	57,597,477	***49,966,684***	***49,972,073***	49,787,449	49,790,835	94,724,569	94,730,191
LOC110257675 (GMCL1L1)	-	-	-	-	***49,845,934***	***49,893,816***	-	-
LOC110257676 (GMCL1L1)	-	-	-	-	49,896,930	49,898,438	-	-
LOC109496524 (ZXDB-like)	57,905,430	57,910,633	***50,250,999***	***50,265,044***	-	-	-	-
ZXDA	57,905,430	57,910,633	-	-	-	-	-	-
LOC105260211 (YBX3-like)	-	-	***50,479,533***	***50,487,761***	-	-	-	-
Human Ampliconic	***58,716,337***	***58,846,090***	-	-	-	-	-	-
Human Ampliconic	***59,467,608***	***59,975,932***	-	-	-	-	-	-
Human Ampliconic	***61,974,100***	***61,984,719***	-	-	-	-	-	-
Human Ampliconic	***63,129,000***	***63,224,416***	-	-	-	-	-	-
Cat Ampliconic	-	-	***52,521,561***	***52,531,618***	-	-	-	-
SPIN4	63,347,228	63,351,339	52,744,389	52,748,367	49,976,956	49,981,312	95,022,507	95,026,682
ARHGEF9	63,634,967	63,786,025	53,234,835	53,436,389	50,159,912	50,578,020	95,048,930	95,196,851
AMER1	64,185,117	64,205,744	53,775,016	53,789,959	50,842,049	50,864,105	95,420,313	95,444,840
ASB12	64,224,192	64,230,631	53,803,025	53,824,360	50,877,158	50,883,739	95,469,407	95,557,993
MTMR8	64,268,081	64,395,453	53,824,313	53,926,316	50,931,199	51,077,796	-	-
LOC101101369 (profilin-1-like)	-	-	53,941,224	53,941,682	51,081,399	51,081,860	-	-
LOC109496915	-	-	54,378,983	54,386,351	-	-	-	-
ZC4H2	64,915,802	65,034,744	54,576,642	54,612,088	51,157,102	51,193,993	95,639,187	95,658,509
ZC3H12B	65,034,815	65,507,887	54,691,632	55,179,184	51,231,526	51,749,627	95,711,558	95,932,656

Table B3.3. Continued

Gene Name	Human Start	Human End	Cat Start	Cat End	Pig Start	Pig End	Mouse Start	Mouse End
LOC109496527 (RPS27-like)	-	-	55,015,556	55,015,836	-	-	-	-
1700010D01Rik	-	-	-	-	-	-	95,732,590	95,733,265
LAS1L	65,512,582	65,534,806	55,183,071	55,202,520	51,753,600	51,773,416	95,920,045	95,956,974
LOC101080617 (RPL37a-like)	-	-	55,302,143	55,302,411	-	-	-	-
MSN	65,588,382	65,741,931	55,364,926	55,430,132	51,844,197	51,967,674	96,096,045	96,168,553
Pig Ampliconic	-	-	-	-	***51,972,371***	***52,070,728***	-	-
VSIG4	66,021,738	66,040,125	55,666,931	55,703,638	52,175,601	52,198,635	96,247,203	96,293,438
Hsf3	-	-	-	-	-	-	96,306,178	96,456,294
HEPH	66,162,526	66,267,389	55,860,077	55,952,047	52,314,911	52,397,384	96,455,436	96,574,485
Gpr165	-	-	-	-	-	-	96,713,468	96,719,388
Pgr15l	-	-	-	-	-	-	97,072,596	97,082,104
EDA2R	66,594,384	66,639,303	56,101,043	56,144,647	52,522,231	52,533,080	97,333,837	97,377,218
Pig Ampliconic	-	-	-	-	***52,611,043***	***52,622,268***	-	-
AR	67,544,032	67,730,619	56,936,744	57,117,143	53,609,113	53,806,778	98,148,757	98,323,218
OPHN1	68,042,344	68,433,841	57,376,752	57,980,658	54,056,504	54,617,045	98,554,280	98,891,414
LOC102899556 (GCSH-like)	-	-	57,802,764	57,808,255	-	-	-	-
YIPF6	68,498,374	68,572,413	58,026,997	58,069,491	54,661,771	54,687,988	98,936,321	98,949,021
STARD8	68,647,669	68,725,842	58,129,992	58,217,287	54,773,189	54,856,716	99,003,228	99,074,728
LOC109496376	-	-	58,287,513	58,292,198	-	-	-	-
EFNB1	68,828,997	68,842,164	58,303,723	58,316,777	54,947,339	54,961,403	99,136,061	99,149,022
LOC105373242	69,031,120	69,084,509	-	-	-	-	-	-
PJA1	69,160,738	69,165,522	58,616,104	58,620,755	55,274,294	55,280,583	99,465,734	99,471,268
FAM155B	69,505,204	69,532,508	58,961,955	58,993,414	55,631,149	55,658,309	99,821,069	99,846,345
EDA	69,616,067	70,039,472	59,075,174	59,521,519	55,770,958	56,103,296	99,975,606	100,400,760
AWAT2	70,040,542	70,049,970	59,522,397	59,531,130	56,101,635	56,108,243	100,402,221	100,494,221

Table B3.3. Continued

Gene Name	Human Start	Human End	Cat Start	Cat End	Pig Start	Pig End	Mouse Start	Mouse End
OTUD6A	70,062,491	70,064,179	59,553,462	59,554,547	56,134,479	56,137,096	100,429,013	100,429,885
IGBP1	70,133,449	70,166,324	59,592,350	59,627,263	56,173,773	56,206,121	100,494,291	100,516,125
DGAT2L6	70,177,483	70,205,703	59,639,474	59,664,645	56,212,494	56,239,641	100,524,838	100,546,108
AWAT1	70,233,489	70,240,661	59,687,850	59,694,126	56,263,082	56,273,169	100,572,247	100,578,206
P2RY4	70,258,166	70,259,804	59,713,784	59,717,735	56,284,736	56,285,833	100,588,358	100,594,927
ARR3	70,268,335	70,281,840	59,728,492	59,743,016	56,295,107	56,310,507	100,605,480	100,618,953
RAB41	70,280,978	70,285,088	-	-	-	-	-	-
PDZD11	70,286,361	70,290,040	59,725,944	59,751,441	56,316,181	56,319,424	100,622,880	100,625,974
KIF4A	70,290,029	70,420,924	59,751,485	59,873,395	56,319,453	56,447,560	100,626,065	100,727,271
GDPD2	70,423,031	70,433,391	59,875,070	59,885,705	56,449,430	56,460,399	100,729,835	100,738,901
LOC105373244	70,427,358	70,436,800	-	-	-	-	-	-
DLG3	70,444,850	70,505,490	59,896,938	59,955,275	56,471,941	56,531,481	100,767,408	100,818,410
TEX11	70,511,227	70,908,717	59,971,709	60,219,061	56,543,933	56,925,104	100,838,648	101,059,639
SLC7A3	70,925,579	70,931,125	60,231,582	60,238,597	56,937,081	56,942,331	101,079,204	101,090,452
SNX12	71,059,247	71,073,426	60,369,727	60,379,917	57,099,521	57,108,636	101,089,367	101,222,583
FOXO4	71,096,149	71,103,534	60,401,229	60,408,532	57,132,549	57,140,377	101,254,259	101,260,873
CXorf65	71,103,865	71,106,973	60,409,014	60,411,307	57,139,975	57,143,070	101,261,373	101,263,992
IL2RG	71,107,404	71,111,631	60,411,656	60,415,380	57,143,568	57,151,242	101,264,385	101,268,255
MED12	71,118,556	71,142,454	60,417,479	60,440,753	57,150,884	57,174,367	101,274,091	101,298,934
NLGN3	71,144,389	71,175,307	60,442,544	60,464,684	57,174,506	57,204,770	101,299,179	101,321,350
GJB1	71,215,212	71,225,215	60,504,288	60,512,002	57,241,990	57,249,496	101,376,376	101,385,629
ZMYM3	71,239,624	71,255,261	60,526,572	60,542,507	57,261,923	57,279,427	101,404,384	101,421,185
NONO	71,283,192	71,301,168	60,556,890	60,572,269	-	-	101,429,651	101,448,593
ITGB1BP2	71,301,275	71,305,371	60,572,392	60,602,903	57,316,490	57,321,614	101,449,109	101,453,541
LOC101093337 (RHOG-like)	-	-	60,610,458	60,619,149	57,368,101	57,380,102	-	-

Table B3.3. Continued

Gene Name	Human Start	Human End	Cat Start	Cat End	Pig Start	Pig End	Mouse Start	Mouse End
TAF1	71,366,220	71,530,525	60,625,287	60,709,079	57,382,211	57,466,700	101,527,572	101,601,789
Pig Ampliconic	-	-	-	-	***57,640,099***	***57,657,993***	-	-
LOC100621166 (CXORF49L)	-	-	-	-	57,678,534	57,682,641	-	-
OGT	71,533,062	71,575,897	60,768,332	60,804,695	57,521,091	57,558,467	101,640,011	101,684,351
ACRC	71,578,024	71,613,583	60,806,504	60,833,925	57,580,582	57,588,938	-	-
CXCR3	71,615,913	71,618,517	60,837,580	60,840,406	57,592,032	57,595,706	101,731,535	101,734,147
CXORF49L	71,667,702	71,671,524	-	-	57,640,600	57,644,707	101,790,332	101,794,456
CXorf49	***71,714,374***	***71,718,285***	60,888,182	60,892,127	57,694,208	57,698,314	101,794,538	101,798,658
CXorf49B	***71,763,290***	***71,767,201***	-	-	-	-	-	-
Gm3858	-	-	-	-	-	-	101,818,196	101,821,816
NHSL2	71,911,073	72,153,286	61,122,714	61,197,506	57,741,172	58,028,266	101,848,669	102,092,055
RGAG4	72,127,110	72,131,901	61,171,431	61,176,259	58,008,195	58,012,965	102,066,542	102,071,378
FLJ44635	72,144,184	72,161,750	-	-	-	-	-	-
PIN4	72,181,676	72,263,964	61,221,092	61,233,441	58,069,758	58,082,243	102,119,465	102,127,673
ERCC6L	72,204,657	72,239,008	61,236,907	61,272,809	58,087,216	58,117,942	102,142,820	102,157,091
RPS4X	72,272,603	72,277,291	61,304,662	61,310,152	58,149,302	58,155,179	102,184,941	102,189,391
CITED1	72,301,638	72,307,187	61,331,226	61,336,743	58,179,084	58,184,778	102,247,378	102,252,181
HDAC8	72,329,516	72,573,103	61,357,126	61,581,292	58,210,182	58,450,681	102,284,638	102,507,710
PHKA1	72,578,814	72,714,181	61,585,672	61,761,695	58,456,097	58,583,216	102,513,975	102,644,286
Gm9112	-	-	-	-	-	-	102,706,888	102,707,670
DMRTC1B	***72,777,170***	***72,848,802***	-	-	-	-	102,707,881	102,715,209
FAM236B	***72,781,865***	***72,782,660***	-	-	-	-	-	-
FAM236D	***72,807,425***	***72,808,819***	-	-	-	-	-	-
DDX46	-	-	61,725,321	61,728,798	-	-	-	-
LOC111558754 (DMRTC1L)	-	-	***61,814,847***	***61,820,535***	***58,696,808***	***58,697,585***	-	-

Table B3.3. Continued

Gene Name	Human Start	Human End	Cat Start	Cat End	Pig Start	Pig End	Mouse Start	Mouse End
LOC102632256 (Dmrtc1-like)	-	-	-	-	-	-	102,764,035	102,779,496
Dmrtc1c1	-	-	-	-	-	-	102,797,880	102,809,950
Dmrtc1c2	-	-	-	-	-	-	102,847,631	102,857,207
1700031F05Rik (Dmrtc1-like)	-	-	-	-	-	-	102,859,166	102,866,353
LOC102632367 (Dmrtc1-like)	-	-	-	-	-	-	102,880,507	102,882,317
DMRTC1	***72,872,025***	***72,943,653***	***61,843,480***	***61,872,443***	***58,697,841***	***58,700,520***	102,903,216	102,934,004
PABPC1L2B (Hominin-sp. duplicate)	73,003,513	73,005,712	-	-	-	-	-	-
ENSSSCG00000031328 (DMRTC1L)	-	-	-	-	***58,717,308***	***58,719,987***	-	-
PABPC1L2B	-	-	61,978,680	61,981,878	***58,830,345***	***58,832,606***	-	-
*FCAB-256L9_Transcript**			62,018,443	62,022,998				
PABPC1L2A	73,077,338	73,079,512	62,086,597	62,087,250	58,851,255	58,852,978	-	-
LOC111558674 (FOXD1-like)	-	-	62,087,568	62,090,480	-	-	-	-
LOC100525624 (NAP1L2L)	-	-	-	-	58,941,164	58,951,122	-	-
Mouse Ampliconic	-	-	-	-	-	-	***103,003,108***	***103,016,033***
4930519F16Rik	-	-	-	-	-	-	103,182,991	103,256,632
NAP1L2	73,212,299	73,214,874	62,149,407	62,151,006	58,969,577	59,176,807	103,184,059	103,186,664
CDX4	73,447,254	73,454,585	-	-	59,089,131	59,100,958	103,321,398	103,330,224
CHIC1	73,563,045	73,687,109	62,267,698	62,348,314	59,176,650	59,230,332	103,356,337	103,409,092
ZCCHC13	74,304,190	74,305,034	62,726,297	62,727,574	59,613,770	59,615,086	103,630,586	103,631,664
SLC16A2	74,421,493	74,533,929	62,866,592	63,014,841	59,748,685	59,861,018	103,697,414	103,821,988
RLIM	74,582,976	74,614,626	63,044,466	63,070,751	59,889,799	59,920,464	103,957,165	103,981,296
NEXMIF	74,732,856	74,925,452	63,128,654	63,290,710	59,967,712	60,013,938	104,076,761	104,201,362
ABCB7	75,053,172	75,156,340	63,428,205	63,596,200	60,186,559	60,313,098	104,280,565	104,413,846
LOC111558772 (PNN-like)	-	-	63,580,928	63,584,015	-	-	-	-
UPRT	75,156,296	75,304,897	63,658,013	63,725,592	60,426,955	60,448,174	104,482,782	104,506,262

Table B3.3. Continued

Gene Name	Human Start	Human End	Cat Start	Cat End	Pig Start	Pig End	Mouse Start	Mouse End
ZDHHC15	75,368,427	75,523,626	63,767,946	64,023,666	60,485,559	60,585,192	104,536,969	104,671,090
MAGEE2	75,782,988	75,785,244	64,102,366	64,104,681	60,707,490	60,711,883	104,854,952	104,857,267
Cat Ampliconic	-	-	***64,216,861***	***64,230,511***	-	-	-	-
PBDC1	76,172,929	76,178,310	64,260,112	64,264,672	60,907,393	60,915,607	105,079,756	105,117,094
MAGEE1	76,427,653	76,431,353	-	-	-	-	105,120,378	105,123,961
Cypt2	-	-	-	-	-	-	105,499,772	105,500,639
Pig Ampliconic	-	-	-	-	***61,474,449***	***61,484,687***	-	-
FGF16	77,447,675	77,456,522	65,150,725	65,159,879	61,551,487	61,561,807	105,764,477	105,776,532
ATRX	77,504,878	77,786,269	65,186,750	65,501,999	61,584,028	61,872,341	105,797,615	105,929,406
LOC108168478 (RPL7A-like)	-	-	-	-	-	-	105,844,465	105,845,490
MAGT1	77,826,364	77,895,568	65,515,652	65,576,466	61,894,237	61,958,511	105,968,084	106,011,899
COX7B	77,899,464	77,905,384	65,582,595	65,588,021	61,962,383	61,969,011	106,015,700	106,022,450
ATP7A	77,910,656	78,050,395	65,592,437	65,768,545	61,972,544	62,110,058	106,027,224	106,128,160
PGAM4	77,967,961	77,969,638	-	-	-	-	-	-
PGK1	78,104,169	78,126,827	65,846,127	65,868,230	62,187,472	62,210,321	106,187,100	106,203,699
TAF9B	78,129,748	78,139,682	65,871,762	65,880,434	62,214,331	62,228,782	106,206,874	106,221,158
FND3C2	-	-	65,901,801	65,961,707	-	-	106,235,246	106,255,372
ERVK-6-like	-	-	65,966,702	65,974,312	-	-	-	-
CYSLTR1	78,271,468	78,327,691	66,073,233	66,122,638	62,380,778	62,407,736	106,576,509	106,603,679
Gm5127	-	-	-	-	-	-	106,583,151	106,710,557
ZCCHC5	78,656,069	78,659,328	66,402,012	66,403,565	62,725,700	62,727,236	106,837,082	106,840,643
LPAR4	78,747,658	78,758,714	66,528,597	66,531,856	62,862,158	62,876,184	106,920,625	106,933,899
P2RY10	78,945,332	78,963,727	66,751,543	66,768,305	63,112,517	63,128,280	107,088,437	107,104,974
P2RY10B	-	-	66,839,861	66,841,078	-	-	107,148,927	107,173,661
GPR174	79,170,972	79,172,229	66,963,517	66,994,590	63,288,948	63,327,621	107,255,878	107,296,771

Table B3.3. Continued

Gene Name	Human Start	Human End	Cat Start	Cat End	Pig Start	Pig End	Mouse Start	Mouse End
ITM2A	79,360,384	79,367,552	67,111,357	67,118,239	63,430,020	63,437,027	107,397,195	107,403,360
TBX22	80,014,730	80,035,828	67,646,961	67,655,352	63,837,576	63,845,704	107,667,964	107,688,980
Cat Ampliconic	-	-	***67,748,352***	***67,765,210***	-	-	-	-
2610002M06Rik	-	-	-	-	-	-	107,782,751	107,816,334
FAM46D	80,335,504	80,445,311	67,981,340	68,066,557	64,120,229	64,122,912	107,792,590	107,872,911
Gm732	-	-	-	-	-	-	107,945,735	107,948,436
BRWD3	80,669,488	80,809,736	68,328,575	68,568,284	64,239,640	64,403,889	108,737,227	108,834,362
HMGN5	81,113,701	81,201,942	66,092,729	66,108,271	64,541,310	64,617,345	109,004,537	109,013,380
SH3BGRL	81,201,804	81,298,547	68,875,480	68,962,945	64,617,350	64,709,006	109,095,407	109,162,467
SH3BGRL3	-	-	68,986,150	68,989,752	-	-	-	-
LOC109496330	-	-	70,625,536	70,683,139	-	-	-	-
Gm6377	-	-	-	-	-	-	109,196,756	109,200,445
Gm34521	-	-	-	-	-	-	109,445,438	109,558,772
Pig Ampliconic	-	-	-	-	***65,080,691***	***65,191,728***	-	-
POU3F4	83,508,261	83,509,767	71,048,729	71,050,280	66,765,044	66,766,642	110,814,379	110,819,108
CYLC1	83,861,126	83,886,700	71,390,149	71,473,379	66,975,125	67,090,165	111,110,418	111,125,961
RPS6KA6	84,058,346	84,187,935	71,571,360	71,766,963	67,234,684	67,415,114	111,387,837	111,538,097
HDX	84,317,874	84,502,578	71,918,729	72,112,986	67,556,764	67,713,744	111,575,204	111,697,079
TEX16	-	-	***72,216,571***	***72,543,138***	-	-	111,896,766	112,127,326
LOC102901599	-	-	72,546,723	72,623,810	-	-	-	-
LOC102899136	-	-	72,624,536	72,629,477	-	-	-	-
4933403O08Rik	-	-	-	-	-	-	112,171,438	112,249,350
APOOL	85,003,852	85,093,317	72,800,927	72,903,870	68,281,781	68,355,009	112,311,352	112,372,755
SATL1	85,092,286	85,116,229	72,921,708	72,938,025	68,356,415	68,383,729	112,384,305	112,406,779
ZNF711	85,243,820	85,273,362	73,118,429	73,145,752	68,475,512	68,499,329	112,600,526	112,635,070

Table B3.3. Continued

Gene Name	Human Start	Human End	Cat Start	Cat End	Pig Start	Pig End	Mouse Start	Mouse End
POF1B	85,277,389	85,379,743	73,149,842	73,272,468	68,503,170	68,579,917	112,638,427	112,698,711
Cat Ampliconic	-	-	***73,282,593***	***73,353,509***	-	-	-	-
Pig Ampliconic	-	-	-	-	***68,586,834***	***68,597,679***	-	-
LOC111558675	-	-	73,612,833	73,617,758	-	-	-	-
CHM	85,861,180	86,047,565	73,736,860	74,003,586	68,873,882	69,066,284	113,040,592	113,185,539
DACH2	86,148,451	86,832,602	74,109,404	74,861,575	69,137,431	69,698,617	113,297,510	113,836,386
KLHL4	87,517,541	87,670,050	-	-	70,334,755	70,436,299	114,474,333	114,561,129
Ube2dnl1	-	-	-	-	-	-	114,905,012	114,905,941
Ube2dnl2	-	-	-	-	-	-	114,907,582	114,908,510
CPXCR1	88,747,225	88,754,785	-	-	71,778,455	71,779,240	116,448,944	116,478,783
LOC110257684 (DPH3 homolog)	-	-	-	-	71,948,753	71,949,100	-	-
Tgif2lx2	-	-	-	-	-	-	118,427,227	118,428,258
TGIF2LX	89,921,941	89,922,883	75,728,003	75,729,287	72,894,181	72,894,630	118,479,708	118,480,729
PABPC5	91,434,598	91,438,584	76,056,725	76,058,934	74,141,707	74,145,941	119,927,196	119,930,165
PCDH11X	91,779,310	92,623,230	76,272,144	76,392,324	74,462,811	75,191,930	120,290,318	120,910,622
NAP1L3	93,670,926	93,673,683	76,889,615	76,891,526	76,147,190	76,149,848	122,394,560	122,397,385
TEN1L	-	-	77,073,467	77,083,883	-	-	-	-
FAM133A	93,673,730	93,712,274	76,892,801	76,943,657	76,149,152	76,181,971	-	-
Cldn34c1	-	-	-	-	-	-	***123,103,523***	***123,145,654***
Astx6	-	-	-	-	-	-	***123,210,420***	***123,213,303***
Srsx	-	-	-	-	-	-	***123,220,985***	***123,222,475***
Gm2411	-	-	-	-	-	-	***123,346,785***	***123,378,718***
Cldn34c2	-	-	-	-	-	-	***123,576,307***	***123,605,024***
Cldn34c3	-	-	-	-	-	-	***123,788,212***	***123,823,149***
Gm5945	-	-	-	-	-	-	***124,016,018***	***124,044,356***

Table B3.3. Continued

Gene Name	Human Start	Human End	Cat Start	Cat End	Pig Start	Pig End	Mouse Start	Mouse End
Vmn2r121	-	-	-	-	-	-	***124,127,339***	***124,135,910***
4932411N23Rik	-	-	-	-	-	-	126,812,462	126,834,004
Gm382	-	-	-	-	-	-	127,039,972	127,063,986
Cldn34c4	-	-	-	-	-	-	127,721,175	127,736,554
DIAPH2	96,684,663	97,600,598	78,669,677	79,681,197	78,602,385	79,520,441	129,749,665	130,465,834
RPA4	96,883,908	96,885,467	-	-	-	-	-	-
LOC105373297 (ERVFC1)	97,841,749	97,849,381	-	-	-	-	-	-
LOC111558778	-	-	79,012,995	79,019,486	-	-	-	-
LOC111558676	-	-	82,027,509	82,032,697	-	-	-	-
LOC100625147 (SNURF)	-	-	-	-	79,811,978	79,812,192	-	-
Pig Ampliconic	-	-	-	-	***80,370,632***	***80,383,729***	-	-
LOC110257759	-	-	-	-	81,543,472	81,545,852	-	-
PCDH19	100,291,644	100,410,273	82,529,430	82,672,588	82,031,365	82,153,159	133,582,860	133,688,993
TNMD	100,584,793	100,599,885	82,822,346	82,837,733	82,269,323	82,292,083	133,851,008	133,865,577
TSPAN6	100,627,108	100,637,104	82,868,565	82,879,598	82,323,634	82,328,919	133,891,068	133,898,429
SRPX2	100,644,166	100,671,299	82,883,433	82,906,770	82,336,741	82,363,868	133,908,416	133,932,448
SYTL4	100,671,964	100,732,148	82,910,136	82,960,747	82,367,546	82,437,439	133,936,385	133,981,822
CSTF2	100,820,359	100,841,520	83,121,207	83,146,786	82,508,739	82,540,861	134,059,138	134,086,822
NOX1	100,843,324	100,874,345	83,147,161	83,171,283	82,540,664	82,575,811	134,086,421	134,137,711
XKRX	100,886,924	100,957,401	83,211,589	83,227,646	82,614,592	82,630,188	134,149,045	134,185,505
ARL13A	100,969,011	100,990,831	83,275,350	83,300,866	82,670,773	82,681,890	134,187,501	134,208,030
TRMT2B	100,973,358	101,052,118	83,303,834	83,334,624	82,638,456	82,741,659	134,222,955	134,277,026
TMEM35A	101,078,847	101,096,367	83,359,846	83,371,855	82,761,552	82,774,952	134,295,225	134,305,969
CENPI	101,098,159	101,181,847	83,374,538	83,453,712	82,776,375	82,826,651	134,308,084	134,363,104
DRP2	101,219,786	101,264,497	83,437,808	83,538,043	82,871,901	82,925,889	134,404,628	134,456,574

Table B3.3. Continued

Gene Name	Human Start	Human End	Cat Start	Cat End	Pig Start	Pig End	Mouse Start	Mouse End
TAF7L	101,268,253	101,293,071	83,541,305	83,559,963	82,930,258	82,951,065	134,460,116	134,476,525
TIMM8A	101,345,656	101,348,969	83,586,166	83,589,010	82,978,128	82,981,038	134,537,256	134,541,670
BTK	101,349,447	101,390,796	83,589,631	83,613,491	82,981,512	83,014,475	134,542,334	134,583,624
RPL36A	101,390,890	101,396,154	83,627,697	83,630,215	83,018,305	83,022,285	134,585,654	134,588,062
GLA	101,397,791	101,408,013	83,631,648	83,639,335	83,024,582	83,032,273	134,588,169	134,601,005
HNRNPH2	101,408,133	101,414,140	83,639,510	83,645,252	83,032,525	83,038,369	134,601,179	134,607,060
ARMCX4	101,418,263	101,535,988	83,701,739	83,713,200	83,092,592	83,130,725	134,685,648	134,697,775
ARMCX1	101,550,531	101,554,700	83,741,474	83,745,659	83,132,775	83,136,924	134,717,913	134,721,917
ARMCX6L	-	-	83,751,767	83,763,976	-	-	-	-
ARMCX6	101,615,118	101,618,001	83,775,462	83,778,400	83,159,108	83,160,331	134,748,455	134,756,666
ARMCX3	101,623,127	101,631,910	83,783,273	83,788,054	83,187,939	83,192,693	134,756,568	134,761,457
ARMCX2	101,655,281	101,659,891	83,820,215	83,824,747	83,213,939	83,218,601	134,804,142	134,809,221
NXF5	101,832,112	101,857,577	83,980,672	84,014,525	83,338,757	83,361,973	134,944,526	134,964,754
ZMAT1	101,882,288	101,932,135	84,075,066	84,119,691	83,396,325	83,559,247	134,971,372	135,009,209
TCEAL2	102,125,688	102,127,712	84,282,478	84,285,901	83,577,181	83,582,946	-	-
Gm15023	-	-	-	-	-	-	***135,167,470***	***135,177,653***
TCEAL6	102,139,961	102,142,416	84,260,200	84,262,241	83,597,173	83,599,216	***135,208,686***	***135,210,687***
BEX5	102,153,707	102,156,014	84,299,324	84,301,493	83,607,366	83,609,797	-	-
NXF2	***102,247,100***	***102,326,720***	84,331,662	84,358,266	-	-	-	-
NXF2B	***102,360,395***	***102,440,008***	-	-	-	-	-	-
Pramel3	-	-	-	-	-	-	***135,302,289***	***135,312,636***
Gm5128	-	-	-	-	-	-	***135,373,250***	***135,383,393***
Gm7903	-	-	-	-	-	-	***135,433,802***	***135,443,997***
AV320801	-	-	-	-	-	-	***135,494,554***	***135,504,748***
Nxf7	-	-	-	-	-	-	135,579,555	135,598,777

Table B3.3. Continued

Gene Name	Human Start	Human End	Cat Start	Cat End	Pig Start	Pig End	Mouse Start	Mouse End
Prame	-	-	-	-	83,674,438	83,677,130	135,612,954	135,627,705
TCP11X2	***102,460,319***	***102,471,812***	84,443,233	84,449,201	83,718,465	83,723,365	135,654,698	135,693,790
TMSB15A	102,513,682	102,516,771	-	-	83,744,536	83,747,599	135,718,667	135,725,261
LOC111556301 (NFX2-like)	-	-	84,512,476	84,514,896	-	-	-	-
ARMCX5	102,599,168	102,604,159	84,537,714	84,539,684	83,746,935	83,769,896	135,742,692	135,747,328
GPRASP1	102,651,366	102,659,083	84,575,130	84,581,481	83,765,210	83,810,090	135,742,692	135,803,468
GPRASP2	102,712,176	102,717,733	84,672,513	84,677,364	83,811,733	83,892,476	135,839,034	135,844,730
BHLHB9	102,720,714	102,752,441	84,679,603	84,692,309	83,895,554	83,911,573	135,841,639	135,891,081
RAB40AL	102,937,272	102,938,300	-	-	-	-	-	-
LOC109496557	-	-	84,736,764	84,738,082	-	-	-	-
LOC109496734 (ENVK7)	-	-	***84,973,344***	***84,977,226***	-	-	-	-
Arxes1	-	-	-	-	-	-	136,033,367	136,034,946
BEX1	103,062,651	103,064,240	***85,146,432***	***85,148,016***	84,029,576	84,031,281	136,066,565	136,068,236
NXF3	103,075,821	103,093,094	***85,080,219***	***85,091,063***	83,980,605	83,989,335	136,072,099	136,085,554
BEX4	103,215,092	103,217,200	91,013,718	91,015,577	84,147,133	84,148,680	136,139,045	136,140,439
LOC105259665 (NFX3L)	-	-	***85,176,176***	***85,188,141***	84,065,215	84,073,826	-	-
LOC110257687 (NXF3L)	-	-	-	-	84,109,452	84,117,997	-	-
TCEAL8	103,252,995	103,255,193	85,300,456	85,302,629	84,175,196	84,177,431	136,168,984	136,172,251
TCEAL5	103,273,690	103,276,869	85,319,830	85,322,845	84,192,480	84,193,013	136,200,948	136,203,851
BEX2	103,309,346	103,311,046	85,335,051	85,336,674	84,223,041	84,224,583	136,213,972	136,215,513
TCEAL7	103,330,186	103,332,326	85,374,791	85,376,841	84,247,651	84,249,858	136,214,914	136,226,100
TCEAL9	103,356,452	103,358,469	85,463,317	85,465,125	84,276,985	84,279,377	136,245,080	136,247,139
BEX3	103,376,323	103,378,164	85,484,264	85,486,102	84,300,645	84,302,124	136,270,253	136,271,978
RAB40A	103,493,266	103,519,489	-	-	-	-	-	-
LOC106508237	-	-	-	-	84,409,372	84,468,158	-	-

Table B3.3. Continued

Gene Name	Human Start	Human End	Cat Start	Cat End	Pig Start	Pig End	Mouse Start	Mouse End
Cat Ampliconic	-	-	***85,758,896***	***85,813,462***	-	-	-	-
TCEAL4	103,576,231	103,587,736	86,271,152	86,273,153	84,475,232	84,477,380	-	-
Kir3dl2	-	-	-	-	-	-	136,447,841	136,544,052
Kir3dl1	-	-	-	-	-	-	136,517,999	136,534,306
TCEAL3	103,607,906	103,609,927	86,279,995	86,282,062	84,485,887	84,490,011	136,590,844	136,668,377
TCEAL1	103,628,720	103,630,953	86,312,017	86,313,994	84,518,999	84,521,020	136,707,969	136,709,873
MORF4L2	103,675,498	103,688,158	86,345,139	86,357,254	84,553,770	84,566,304	136,732,942	136,743,686
GLRA4	103,707,344	103,728,624	86,372,981	86,398,209	84,584,907	84,612,860	136,756,628	136,783,517
TMEM31	103,710,909	103,714,032	86,381,429	86,382,937	-	-	-	-
PLP1	103,776,506	103,792,619	86,445,195	86,461,282	84,676,627	84,683,517	136,822,166	136,838,582
RAB9B	103,822,325	103,832,282	86,493,535	86,502,854	84,721,269	84,733,000	136,858,151	136,868,540
Pig Ampliconic	-	-	-	-	***84,779,597***	***84,806,994***	-	-
TMSB15B	***103,962,627***	***103,965,990***	86,571,122	86,572,647	84,815,565	84,871,809	136,954,988	136,976,874
Tmsb15b2	-	-	-	-	-	-	136,954,988	136,957,979
H2BFWT	104,011,147	104,013,687	-	-	-	-	-	-
H2BFM	104,039,753	104,042,454	86,551,551	86,551,817	-	-	136,927,325	136,928,373
Tmsb15b1	-	-	-	-	-	-	136,974,022	136,976,874
SLC25A53	104,099,214	104,157,027	86,596,574	86,606,525	84,860,180	84,871,866	136,981,116	137,038,302
ZCCHC18	104,112,416	104,121,330	86,606,693	86,612,694	84,870,668	84,875,040	136,992,379	136,996,923
FAM199X	104,157,074	104,195,902	86,666,970	86,701,639	84,937,153	84,966,438	137,049,594	137,082,503
ESX1	104,250,038	104,254,918	86,757,059	86,762,044	84,995,241	85,007,557	137,115,397	137,122,131
LOC110257688 (CFAP251-like)	-	-	-	-	85,039,485	85,040,165	-	-
IL1RAPL2	104,566,315	105,769,729	87,075,175	88,285,364	85,347,867	86,464,670	137,570,754	138,849,077
TEX13A	105,218,929	105,220,693	87,724,903	87,726,548	85,955,144	85,956,458	138,208,158	138,209,591
NRK	105,821,944	105,958,610	88,335,362	88,480,960	86,534,467	86,663,340	138,914,318	139,010,532

Table B3.3. Continued

Gene Name	Human Start	Human End	Cat Start	Cat End	Pig Start	Pig End	Mouse Start	Mouse End
SERPINA7	106,032,439	106,038,727	88,564,912	88,568,711	86,708,624	86,714,324	139,079,249	139,085,270
4930513O06Rik	-	-	-	-	-	-	139,086,243	139,093,403
4933428M09Rik	-	-	-	-	-	-	139,179,215	139,181,009
MUM1L1	106,168,278	106,208,961	***88,629,709***	***88,656,594***	86,785,317	86,814,631	139,210,030	139,238,335
Human Ampliconic	***106,266,280***	***106,278,816***	-	-	-	-	-	-
LOC100515519 (NCL-like)	-	-	-	-	86,974,106	86,976,917	-	-
LOC101101037 (SUMO2)	-	-	88,934,371	89,096,038	-	-	-	-
Trap1a	-	-	-	-	-	-	139,333,683	139,338,169
CXorf57	106,611,930	106,679,443	89,108,893	89,194,421	87,025,603	87,103,503	139,480,226	139,554,589
RNF128	106,693,838	106,797,016	89,207,826	89,316,944	87,111,875	87,211,605	139,563,340	139,673,145
TBC1D8B	106,802,689	106,876,150	89,322,911	89,382,760	87,217,327	87,291,054	139,684,960	139,754,221
RIPPLY1	106,900,063	106,903,331	89,413,404	89,416,691	87,315,355	87,319,580	139,779,681	139,782,353
CLDN2	106,900,164	106,930,861	89,432,933	89,443,334	87,352,695	87,355,724	139,800,808	139,811,386
MORC4	106,940,300	107,000,244	89,454,771	89,505,217	87,369,311	87,427,296	139,821,635	139,871,691
RBM41	107,061,885	107,118,827	89,620,911	89,683,135	87,512,887	87,575,118	139,889,511	139,998,595
NUP62CL	107,123,427	107,206,440	89,687,468	89,768,043	87,578,656	87,658,296	140,006,802	140,062,707
PIH1D3	107,206,629	107,244,243	89,768,189	89,804,097	87,658,355	87,696,318	140,062,879	140,106,797
LOC101085786 (RPS27-like)	-	-	89,933,289	89,933,540	-	-	-	-
LOC111558744 (WH1-like)	-	-	89,988,732	89,989,977	-	-	-	-
FRMPD3	107,522,450	107,605,264	90,040,012	90,122,927	87,980,635	88,053,070	140,277,679	140,435,467
PRPS1	107,628,424	107,651,026	90,136,326	90,157,906	88,074,861	88,101,925	140,456,603	140,476,140
TSC22D3	107,713,221	107,777,329	90,226,095	90,288,609	88,171,325	88,233,583	140,539,528	140,600,522
MID2	107,820,961	107,931,637	90,307,743	90,442,685	88,294,032	88,394,885	140,664,304	140,767,715
LOC105373310	107,968,531	107,986,048	90,468,689	90,494,958	88,403,475	88,466,669	-	-
Eif2c5	-	-	-	-	-	-	140,776,288	140,796,880

Table B3.3. Continued

Gene Name	Human Start	Human End	Cat Start	Cat End	Pig Start	Pig End	Mouse Start	Mouse End
TEX13B	107,980,864	107,982,370	90,479,020	90,481,916	88,431,789	88,434,544	140,808,307	140,813,433
VSIG1	108,018,886	108,079,184	90,542,058	90,579,781	88,468,498	88,534,958	140,907,515	140,939,472
PSMD10	108,084,205	108,091,644	90,586,030	90,594,129	88,540,261	88,548,593	140,948,425	140,956,711
ATG4A	108,088,900	108,154,671	90,594,280	90,651,145	88,545,924	88,633,328	140,953,893	141,164,270
COL4A6	108,155,607	108,439,497	90,652,489	90,968,064	88,634,925	88,750,938	141,165,403	141,474,076
COL4A5	108,439,844	108,697,545	90,969,508	91,209,824	88,958,837	89,183,306	141,475,370	141,689,236
IRS4	108,719,949	108,736,584	91,231,558	91,249,803	89,204,348	89,221,463	141,710,998	141,725,217
GUCY2F	109,372,061	109,482,056	91,654,427	91,766,964	89,557,952	89,653,491	142,079,289	142,197,276
NXT2	109,535,781	109,544,698	91,797,801	91,804,611	89,684,733	89,691,808	142,226,770	142,239,700
KCNE5	109,623,700	109,625,164	91,862,103	91,863,562	89,741,187	89,742,732	142,304,753	142,306,198
ACSL4	109,641,330	109,733,403	91,886,462	91,945,715	89,760,863	89,838,103	142,317,993	142,390,800
LOC102902491	-	-	92,055,713	92,058,745	-	-	-	-
TMEM164	110,002,413	110,184,438	92,232,286	92,397,871	90,105,293	90,289,674	142,680,347	142,843,494
LOC101928589	110,175,773	110,177,788	-	-	-	-	-	-
AMMECR1	110,194,186	110,440,233	92,414,850	92,526,228	90,304,918	90,420,207	142,853,474	142,966,728
RGAG1	110,358,101	110,456,334	92,646,345	92,657,590	90,553,099	90,565,001	143,057,943	143,105,330
CHRDL1	110,673,856	110,796,058	92,802,545	92,920,974	90,703,422	90,827,216	143,285,674	143,394,265
PAK3	110,944,241	111,227,361	93,049,536	93,308,221	90,941,976	91,244,331	143,518,371	143,797,796
CAPN6	111,245,099	111,270,546	93,321,132	93,344,989	91,258,696	91,288,128	143,802,236	143,827,412
DCX	111,293,779	111,412,232	93,375,907	93,486,009	91,317,671	91,644,572	143,855,842	143,933,275
SERTM2	111,511,662	111,522,399	93,594,114	93,605,579	91,534,775	91,549,167	144,153,695	144,166,274
ALG13	111,680,741	111,764,164	93,720,536	93,782,588	91,644,586	91,708,695	144,317,966	144,374,450
TRPC5	111,774,314	112,082,943	93,794,525	94,068,124	91,716,334	91,988,364	144,377,327	144,688,180
TRPC5OS	111,876,051	111,903,990	93,819,527	93,897,235	91,821,128	91,823,780	144,455,781	144,477,063
RTL4	112,083,025	112,457,245	94,405,940	94,419,511	91,920,451	92,333,326	144,688,218	145,122,414

Table B3.3. Continued

Gene Name	Human Start	Human End	Cat Start	Cat End	Pig Start	Pig End	Mouse Start	Mouse End
LOC100738042 (MIC10-like)	-	-	-	-	92,197,909	92,199,862	-	-
LHFPL1	112,630,651	112,680,178	94,529,557	94,569,408	92,422,913	92,542,291	145,290,359	145,372,308
AMOT	112,774,877	112,840,946	94,651,745	94,715,642	92,581,995	92,647,872	145,446,424	145,505,533
HTR2C	114,584,078	114,910,061	96,129,986	96,431,438	94,053,415	94,313,352	146,962,513	147,197,277
PERK10	-	-	96,453,223	96,454,335	-	-	-	-
LOC111556295	-	-	96,454,338	96,459,055	-	-	-	-
IL13RA2	115,003,975	115,017,644	96,523,327	96,570,455	94,402,665	94,455,609	147,383,478	147,429,192
LRCH2	115,110,611	115,234,082	96,625,663	96,726,637	94,486,560	94,589,930	147,470,375	147,554,102
RBMXL3	115,189,400	115,192,868	-	-	-	-	-	-
LOC102900824 (SET-like)	-	-	96,673,478	96,674,889	-	-	-	-
GM15128	-	-	-	-	-	-	***147,789,653***	***147,833,116***
GM15080	-	-	-	-	-	-	***147,992,993***	***148,033,174***
GM15107	-	-	-	-	-	-	***148,180,675***	***148,224,710***
GM15108	-	-	-	-	-	-	***148,488,949***	***148,521,446***
GM15127	-	-	-	-	-	-	***148,663,588***	***148,703,849***
LUZP4	115,289,717	115,307,556	96,803,935	96,842,171	94,631,490	94,647,712	***148,857,694***	***148,924,139***
LOC105373314	115,454,266	115,465,553	-	-	-	-	-	-
PLS3	115,560,850	115,650,861	97,062,586	97,156,737	94,885,098	94,971,664	75,785,654	75,875,178
Human Ampliconic	***115,843,208***	***115,886,139***	-	-	-	-	-	-
Pig Ampliconic	-	-	-	-	***95,013,653***	***95,139,430***	-	-
AGTR2	116,170,705	116,174,972	97,342,561	97,346,822	95,267,709	95,272,237	21,484,549	21,488,833
SLC6A14	116,436,579	116,461,458	97,554,010	97,577,942	95,482,494	95,507,200	21,714,900	21,742,358
CT83	116,461,686	116,463,033	97,579,892	97,581,667	95,508,720	95,510,921	-	-
KLHL13	117,897,813	118,117,340	98,678,542	98,882,383	96,545,099	96,950,242	23,219,271	23,365,108
LOC110257690	-	-	-	-	95,554,791	95,654,903	-	-

Table B3.3. Continued

Gene Name	Human Start	Human End	Cat Start	Cat End	Pig Start	Pig End	Mouse Start	Mouse End
LOC414400 (RPLP1L)	-	-	-	-	96,800,950	96,875,862	-	-
WDR44	118,345,997	118,449,961	99,101,847	99,181,312	96,950,288	97,027,069	23,692,952	23,806,026
Gm4907	-	-	-	-	-	-	23,882,524	23,907,602
Gm5934	-	-	-	-	-	-	24,474,308	24,499,163
Gm4297	-	-	-	-	-	-	24,552,250	24,573,305
Gm5935	-	-	-	-	-	-	24,753,162	24,775,164
Gm5169	-	-	-	-	-	-	***25,277,487***	***25,301,455***
Gm1993	-	-	-	-	-	-	25,548,501	25,570,449
Gm5168	-	-	-	-	-	-	***26,028,214***	***26,051,041***
Gm2012	-	-	-	-	-	-	***26,200,882***	***26,231,323***
Gm2030	-	-	-	-	-	-	***26,287,199***	***26,310,044***
Gm2036	-	-	-	-	-	-	***26,414,963***	***26,415,310***
Slx	-	-	-	-	-	-	***26,522,657***	***26,545,565***
Gm14525	-	-	-	-	-	-	***26,672,780***	***26,702,638***
Gm6121	-	-	-	-	-	-	***26,912,743***	***26,935,669***
Gm10230	-	-	-	-	-	-	***27,472,033***	***27,494,939***
Gm10058	-	-	-	-	-	-	***27,924,021***	***27,946,885***
Gm4836	-	-	-	-	-	-	***28,270,710***	***28,293,632***
Gm10147	-	-	-	-	-	-	***28,584,995***	***28,607,874***
Gm10096	-	-	-	-	-	-	***28,931,807***	***28,954,702***
Gm10488	-	-	-	-	-	-	***29,976,314***	***29,999,162***
Gm14632	-	-	-	-	-	-	***30,329,740***	***30,352,646***
Gm10487	-	-	-	-	-	-	***30,820,542***	***30,843,469***
Btbd35f29	-	-	-	-	-	-	***31,117,576***	***31,119,415***
LOC108168466 (SPIN2A-like)	-	-	-	-	-	-	***31,254,636***	***31,255,381***

Table B3.3. Continued

Gene Name	Human Start	Human End	Cat Start	Cat End	Pig Start	Pig End	Mouse Start	Mouse End
Btbd35f26	-	-	-	-	-	-	***31,336,219***	***31,338,442***
Btbd35f19	-	-	-	-	-	-	***31,381,876***	***31,384,047***
Spin2e	-	-	-	-	-	-	***31,618,191***	***31,619,833***
Btbd35f25	-	-	-	-	-	-	***31,715,299***	***31,717,235***
Gm21637	-	-	-	-	-	-	***31,960,926***	***31,962,821***
Btbd35f22	-	-	-	-	-	-	***32,047,582***	***32,049,551***
Gm5923	-	-	-	-	-	-	***32,158,701***	***32,160,371***
Btbd35f2	-	-	-	-	-	-	***32,278,394***	***32,280,342***
Gm21466	-	-	-	-	-	-	***32,437,174***	***32,438,844***
Btbd35f1	-	-	-	-	-	-	***32,560,055***	***32,562,009***
Gm5926	-	-	-	-	-	-	***32,649,778***	***32,651,426***
Btbd35f21	-	-	-	-	-	-	***32,726,137***	***32,728,057***
Btbd35f30	-	-	-	-	-	-	***32,894,527***	***32,896,366***
Btbd35f5	-	-	-	-	-	-	***32,973,897***	***32,975,846***
Btbd35f12	-	-	-	-	-	-	***33,313,338***	***33,315,288***
Gm2854	-	-	-	-	-	-	***33,398,517***	***33,400,130***
LOC100862179 (SPIN2A-like)	-	-	-	-	-	-	***33,476,982***	***33,478,575***
Btbd35f15	-	-	-	-	-	-	***33,575,387***	***33,577,346***
Btbd35f13	-	-	-	-	-	-	***33,657,140***	***33,659,082***
Btbd35f14	-	-	-	-	-	-	***33,700,649***	***33,702,349***
Gm46683	-	-	-	-	-	-	***33,777,828***	***33,778,538***
Btbd35f6	-	-	-	-	-	-	***33,954,615***	***33,956,624***
Btbd35f9	-	-	-	-	-	-	***34,159,642***	***34,161,734***
LOC108168468 (SPIN2A-like)	-	-	-	-	-	-	***34,237,100***	***34,238,592***
Btbd35f8	-	-	-	-	-	-	***34,450,295***	***34,452,516***

Table B3.3. Continued

Gene Name	Human Start	Human End	Cat Start	Cat End	Pig Start	Pig End	Mouse Start	Mouse End
Gm10486	-	-	-	-	-	-	***35,067,384***	***35,090,316***
Gm14819	-	-	-	-	-	-	***35,404,215***	***35,427,094***
DOCK11	118,495,664	118,686,162	99,233,868	99,435,031	97,068,488	97,267,528	35,888,621	36,076,570
IL13RA1	118,726,954	118,794,533	99,469,068	99,526,169	97,300,960	97,353,754	36,112,108	36,171,262
ZCCHC12	118,823,743	118,826,973	99,554,064	99,557,311	97,379,908	97,384,384	36,195,904	36,199,158
LONRF3	118,974,614	119,022,925	99,660,309	99,699,983	97,494,305	97,547,813	36,328,358	36,366,857
LOC111558681	-	-	99,739,151	99,740,665	-	-	-	-
LOC105259672 (CT47A-like)	-	-	99,740,979	99,743,939	-	-	-	-
Gm33383 (Ct47b-like)	-	-	-	-	-	-	36,401,263	36,403,446
Gm6268 (Ct47-like)	-	-	-		-	-	36,403,477	36,404,654
KIAA1210	119,078,635	119,151,082	99,747,609	99,806,296	97,581,400	97,651,765	36,422,381	36,513,829
PGRMC1	119,236,245	119,244,466	99,892,970	99,901,650	97,735,402	97,743,424	36,598,193	36,606,079
AKAP17B	119,244,740	119,273,600	99,904,929	99,934,875	97,745,976	97,873,665	36,608,173	36,645,472
SLC25A43	119,399,295	119,454,478	100,039,630	100,072,676	97,868,804	97,906,771	36,743,549	36,777,307
SLC25A5	119,468,400	119,471,396	100,083,227	100,086,380	97,919,295	97,922,245	36,795,597	36,798,808
CXorf56	119,538,149	119,565,434	100,112,840	100,167,428	97,974,596	97,986,281	36,827,551	36,864,303
UBE2A	119,574,467	119,584,429	100,177,117	100,186,714	97,925,136	98,007,551	36,874,295	36,884,222
NKRF	119,588,337	119,606,464	100,189,350	100,205,626	98,008,870	98,025,423	36,886,440	36,903,344
SEPT6	119,615,724	119,693,370	100,212,884	100,280,713	98,033,212	98,106,082	36,911,272	36,990,050
SOWAHD	119,758,613	119,760,202	100,354,892	100,357,716	98,154,437	98,158,808	37,048,845	37,050,418
RPL39	119,786,504	119,791,659	100,382,312	100,386,066	98,178,148	98,181,439	37,082,518	37,085,222
UPF3B	119,805,311	119,853,028	100,402,201	100,418,103	98,184,646	98,206,135	37,091,678	37,110,400
RNF113A	119,870,532	119,871,828	100,434,695	100,435,906	98,218,689	98,219,931	37,191,222	37,192,467
NDUFA1	119,871,771	119,876,666	100,435,999	100,439,954	98,219,717	98,223,873	37,187,588	37,191,238
AKAP14	119,895,973	119,920,716	100,449,079	100,469,552	98,234,469	98,255,396	37,150,698	37,168,842

Table B3.3. Continued

Gene Name	Human Start	Human End	Cat Start	Cat End	Pig Start	Pig End	Mouse Start	Mouse End
NKAP	119,925,050	119,943,772	100,469,501	100,496,661	98,233,504	98,276,839	37,126,763	37,150,746
RHOXF2B	120,072,264	120,077,742	-	-	-	-	-	-
RHOXF1	***120,109,049***	***120,117,651***	100,525,559	100,528,519	98,304,777	98,310,448	-	-
RHOXF2	***120,158,534***	***120,164,039***	-	-	-	-	-	-
LOC111558775 (Rhox2b-like)	-	-	100,589,354	100,598,260	-	-	-	-
LOC101093839 (RHOXF1)	-	-	100,643,391	100,676,554	-	-	-	-
LOC102902213 (ESX1-like)	-	-	100,670,140	100,683,604	98,336,502	98,366,377	-	-
Gm9	-	-	-	-	-	-	37,208,502	37,211,041
Rhox1	-	-	-	-	-	-	37,213,804	37,222,258
Rhox2a	-	-	-	-	-	-	37,244,992	37,249,690
Rhox3a	-	-	-	-	-	-	***37,249,919***	***37,258,978***
Rhox4a	-	-	-	-	-	-	***37,265,067***	***37,270,136***
Rhox3a2	-	-	-	-	-	-	***37,380,383***	***37,384,431***
Rhox4a2	-	-	-	-	-	-	***37,390,860***	***37,395,621***
Rhox2b	-	-	-	-	-	-	37,411,736	37,416,844
Rhox4b	-	-	-	-	-	-	37,432,494	37,437,278
Rhox2c	-	-	-	-	-	-	37,453,480	37,458,417
Rhox3c	-	-	-	-	-	-	37,469,868	37,473,961
Rhox4c	-	-	-	-	-	-	37,480,337	37,485,124
Rhox2d	-	-	-	-	-	-	37,493,410	37,497,715
Rhox4d	-	-	-	-	-	-	37,514,385	37,519,176
Rhox2e	-	-	-	-	-	-	37,530,322	37,541,851
Rhox3e	-	-	-	-	-	-	37,546,975	37,550,897
Rhox4e	-	-	-	-	-	-	37,557,412	37,562,283
Rhox2f	-	-	-	-	-	-	37,571,421	37,576,159

Table B3.3. Continued

Gene Name	Human Start	Human End	Cat Start	Cat End	Pig Start	Pig End	Mouse Start	Mouse End
Rhox3f	-	-	-	-	-	-	37,581,352	37,585,496
Rhox4f	-	-	-	-	-	-	37,602,895	37,607,686
Rhox3g	-	-	-	-	-	-	37,623,431	37,628,139
Rhox2g	-	-	-	-	-	-	37,639,112	37,643,470
Rhox4g	-	-	-	-	-	-	37,646,501	37,651,327
Rhox3h	-	-	-	-	-	-	37,657,688	37,668,764
Rhox2h	-	-	-	-	-	-	37,668,997	37,673,277
Rhox5	-	-	-	-	-	-	37,754,608	37,808,878
Rhox6	-	-	-	-	-	-	37,827,055	37,829,857
Rhox7a	-	-	-	-	-	-	37,831,686	37,841,171
Rhox8	-	-	-	-	-	-	37,874,776	37,890,766
Rhox9	-	-	-	-	-	-	37,899,097	37,901,770
LOC102637882	-	-	-	-	-	-	37,940,614	37,945,671
Rhox10	-	-	-	-	-	-	38,066,475	38,071,688
Rhox11	-	-	-	-	-	-	38,076,598	38,085,139
Rhox12	-	-	-	-	-	-	38,104,062	38,110,907
Rhox13	-	-	-	-	-	-	38,120,840	38,129,966
ZBTB33	120,250,752	120,258,396	100,802,088	100,809,770	98,370,137	98,413,686	38,189,793	38,197,046
TMEM255A	120,250,984	120,311,556	100,810,623	100,867,057	98,412,286	98,464,683	38,196,573	38,252,490
ATP1B4	120,362,085	120,383,249	100,915,971	100,937,871	98,514,286	98,535,154	38,316,122	38,336,784
LOC109496592 (TEX13C)	-	-	100,959,844	100,963,170	98,565,476	98,570,266	-	-
ADH4	-	-	-	-	98,582,002	98,583,002	-	-
LAMP2	120,426,148	120,469,349	100,975,929	101,012,517	98,586,718	98,623,939	38,401,357	38,456,463
CUL4B	120,524,589	120,575,829	101,076,321	101,132,937	98,668,945	98,730,582	38,531,616	38,576,196
MCTS1	120,603,889	120,621,159	101,156,495	101,170,352	98,729,858	98,742,967	38,600,444	38,613,524

Table B3.3. Continued

Gene Name	Human Start	Human End	Cat Start	Cat End	Pig Start	Pig End	Mouse Start	Mouse End
C1GALT1C1	120,625,674	120,630,150	101,180,796	101,185,387	98,752,711	98,758,428	38,630,783	38,635,143
6030498E09Rik	-	-	-	-	-	-	38,772,780	38,962,688
Cypt15	-	-	-	-	-	-	39,346,266	39,346,963
Cypt14	-	-	-	-	-	-	39,862,919	39,863,604
CT47B1	120,872,554	120,875,925	-	-	-	-	-	-
CT47A12	***120,877,490***	***120,932,399***	-	-	-	-	-	-
CT47A11	***120,933,840***	***120,937,260***	-	-	-	-	-	-
CT47A10	***120,938,701***	***120,942,121***	-	-	-	-	-	-
CT47A9	***120,943,561***	***120,946,981***	-	-	-	-	-	-
CT47A8	***120,948,422***	***120,951,842***	-	-	-	-	-	-
CT47A7	***120,953,282***	***120,956,600***	-	-	-	-	-	-
CT47A6	***120,958,165***	***120,961,588***	-	-	-	-	-	-
CT47A5	***120,963,026***	***120,966,446***	-	-	-	-	-	-
CT47A4	***120,967,886***	***120,971,306***	-	-	-	-	-	-
CT47A3	***120,972,746***	***120,976,166***	-	-	-	-	-	-
CT47A2	***120,977,606***	***120,981,026***	-	-	-	-	-	-
CT47A1	***120,982,466***	***120,985,886***	-	-	98,978,636	98,984,740	-	-
GLUD2	121,047,608	121,049,942	-	-	-	-	-	-
GRIA3	123,184,243	123,490,915	103,173,624	103,439,996	100,827,293	101,109,213	41,401,118	41,678,601
THOC2	123,600,561	123,733,054	103,528,947	103,642,524	101,170,969	101,284,623	41,794,992	41,911,919
XIAP	123,859,812	123,913,979	103,741,876	103,788,765	101,399,374	101,440,021	42,059,615	42,109,664
STAG2	123,960,560	124,102,656	103,832,748	103,972,138	101,476,117	101,617,405	42,149,319	42,277,186
TEX13D	124,332,660	124,336,862	104,060,755	104,156,331	101,806,861	101,810,876	42,489,361	42,491,953
SH2D1A	124,346,282	124,373,160	104,166,566	104,188,315	101,824,602	101,847,265	42,502,565	42,522,097
TENM1	124,374,680	125,204,305	104,192,322	104,980,011	101,853,340	102,641,299	42,527,866	43,429,155

Table B3.3. Continued

Gene Name	Human Start	Human End	Cat Start	Cat End	Pig Start	Pig End	Mouse Start	Mouse End
TEX13C	125,319,871	125,325,214	105,110,553	105,114,826	-	-	43,590,989	43,593,530
LOC111558717 (SLC25A5-like)	-	-	105,479,849	105,482,968	-	-	-	-
DCAF12L2	126,163,499	126,166,097	105,810,090	105,813,296	103,322,468	103,327,458	44,365,457	44,368,337
DCAF12L1	126,549,383	126,552,859	106,146,978	106,148,750	103,369,987	103,641,354	44,786,567	44,790,197
LOC101098936 (COA1 homolog)	-	-	106,332,128	106,332,481	-	-	-	-
PRR32	126,819,764	126,821,786	106,379,225	106,381,245	103,813,092	103,815,208	45,090,904	45,092,791
ACTRT1	128,050,962	128,052,403	107,750,209	107,751,312	104,936,448	104,941,473	46,329,007	46,330,345
SMARCA1	129,446,501	129,523,564	109,008,428	109,080,923	106,039,961	106,116,156	47,809,370	47,892,575
OCRL	129,532,737	129,592,561	109,090,403	109,148,226	106,124,991	106,188,066	47,912,401	47,966,428
APLN	129,645,259	129,654,956	109,195,678	109,205,006	106,236,746	106,246,563	48,025,146	48,034,852
XPNPEP2	129,738,970	129,769,549	109,281,773	109,308,423	106,333,785	106,363,412	48,106,913	48,136,982
SASH3	129,779,916	129,795,201	109,317,178	109,330,795	106,373,871	106,389,454	48,146,436	48,161,565
ZDHHC9	129,803,288	129,843,934	109,340,044	109,371,945	106,397,139	106,436,103	48,171,967	48,208,702
CCDC169	-	-	-	-	106,457,036	106,457,680	-	-
UTP14A	129,906,121	129,929,762	109,414,561	109,433,937	106,479,028	106,500,389	48,256,934	48,282,450
BCORL1	129,980,285	130,058,083	109,471,626	109,535,519	106,547,754	106,613,130	48,341,358	48,408,068
ELF4	130,064,920	130,110,713	109,539,955	109,590,032	106,617,747	106,661,340	48,411,049	48,463,132
RAB33A	130,110,633	130,184,872	109,636,105	109,646,903	106,708,402	106,723,803	48,519,285	48,530,240
AIFM1	130,129,362	130,165,887	109,598,065	109,631,160	106,676,596	106,708,290	48,474,944	48,513,563
ZNF280C	130,202,699	130,268,948	109,638,863	109,724,019	106,735,129	106,795,789	48,541,626	48,594,648
SLC25A14	130,339,888	130,373,361	109,765,355	109,794,550	106,828,695	106,859,128	48,623,418	48,662,298
GPR119	130,384,345	130,385,537	109,807,236	109,814,034	106,872,225	106,873,299	48,667,979	48,674,478
RBMX2	130,401,969	130,413,343	109,826,582	109,837,672	106,880,912	106,890,391	48,695,004	48,710,719
LOC101091264 (FSIP2L)	-	-	109,847,671	109,927,335	106,902,342	106,978,056	-	-
ENOX2	130,622,330	130,903,317	110,035,277	110,302,632	107,084,378	107,368,520	49,009,707	49,288,244

Table B3.3. Continued

Gene Name	Human Start	Human End	Cat Start	Cat End	Pig Start	Pig End	Mouse Start	Mouse End
ARHGAP36	131,058,242	131,089,885	110,428,203	110,480,926	107,368,552	107,537,747	49,448,359	49,500,250
Olfr1320	-	-	-	-	-	-	49,683,504	49,684,463
Olfr1321	-	-	-	-	-	-	49,726,973	49,727,932
IGSF1	131,273,506	131,289,457	110,604,629	110,620,988	107,674,966	108,050,736	49,782,536	49,797,907
LOC101091768 (OR10A4L)	-	-	110,675,889	110,676,794	107,736,747	107,737,673	49,885,469	49,886,401
LOC100511059 (OR1020L)	-	-	-	-	107,752,365	107,762,735	-	-
LOC101092250 (OR1052L)	-	-	110,755,004	110,755,927	107,820,069	107,821,040	50,009,306	50,010,235
LOC101092501 (OR1468L)	-	-	110,777,253	110,778,318	107,894,520	107,895,892	50,425,496	50,426,494
OR13H1	131,544,074	131,545,000	110,819,934	110,820,860	107,908,613	107,912,960	-	-
LOC101093004 (OR13H1L)	-	-	110,884,505	110,893,925	107,949,245	107,956,580	-	-
STK26	132,023,217	132,075,943	111,143,347	111,197,945	108,179,230	108,237,669	50,841,174	50,893,103
FRMD7	132,076,986	132,128,022	111,199,495	111,251,536	108,241,116	108,294,188	50,892,645	50,943,619
RAP2C	132,203,024	132,219,480	111,298,726	111,314,910	108,332,623	108,348,432	51,003,912	51,020,092
MBNL3	132,369,314	132,490,030	111,421,653	111,533,165	108,488,224	108,612,896	51,113,494	51,206,035
HS6ST2	132,626,010	132,961,395	111,672,926	111,964,671	108,755,577	109,043,976	51,386,637	51,681,705
USP26	133,024,631	133,083,768	112,018,802	112,072,097	109,092,622	109,095,548	51,752,667	51,801,233
TFDP3	133,216,669	133,218,348	-	-	-	-	-	-
1700080O16Rik (Magea10L)	-	-	-	-	-	-	51,968,693	51,972,864
GPC4	133,301,036	133,415,177	112,244,732	112,354,114	109,306,489	109,421,947	52,050,885	52,164,923
GPC3	133,535,745	133,985,646	112,443,529	112,865,781	109,536,495	109,973,605	52,272,426	52,613,974
LOC109496610 (COXB1-like)	-	-	112,978,811	112,979,341	-	-	-	-
CCDC160	134,237,047	134,246,207	113,041,785	113,050,638	110,159,833	110,170,354	52,791,200	52,799,474
PHF6	134,373,312	134,428,792	113,134,108	113,180,666	110,249,010	110,311,503	52,912,214	52,956,961
HPRT1	134,460,145	134,500,668	113,204,772	113,240,012	110,321,144	110,357,902	52,988,078	53,021,660
LOC102900338 (RPS3A-like)	-	-	113,220,613	113,221,812	-	-	-	-

Table B3.3. Continued

Gene Name	Human Start	Human End	Cat Start	Cat End	Pig Start	Pig End	Mouse Start	Mouse End
PLAC1	134,565,838	134,764,322	113,288,573	113,289,425	110,403,257	110,592,117	53,069,993	53,240,122
FAM122B	134,769,566	134,797,232	113,289,369	113,492,901	110,570,091	110,617,779	53,243,415	53,269,852
FAM122C	134,796,413	134,855,137	113,498,257	113,599,759	110,619,227	110,682,249	53,273,419	53,331,186
MOSPD1	134,887,626	134,915,344	113,574,739	113,598,049	110,693,538	110,786,323	53,344,594	53,370,562
SMIM10	134,990,938	134,992,473	113,664,993	113,666,004	110,754,797	110,761,398	-	-
FAM127C	135,020,504	135,022,536	-	-	110,793,198	110,794,496	53,607,922	53,609,135
FAM127A	135,032,303	135,033,546	-	-	110,810,150	110,811,475	53,642,488	53,643,763
LOC110257712 (FAM127L)	-	-	-	-	110,813,912	110,815,239	-	-
FAM127B	135,050,932	135,052,191	-	-	110,833,247	110,834,588	53,669,177	53,670,408
SMIM10L2B	135,094,985	135,098,634	-	-	-	-	-	-
FCAB-89F20_Transcript1	-	-	113,531,770*	113,689,137*	-	-	-	-
CT55	***135,156,536***	***135,171,827***	113,827,541	113,851,995	110,848,227	110,890,277	53,724,831	53,738,565
LOC110255250 (SIM10L2A)	-	-	-	-	110,872,816	110,876,723	-	-
ZNF75D	***135,248,589***	***135,344,109***	113,895,482	113,909,163	110,890,310	110,904,067	-	-
LOC101928677 (ETDAL)	***135,252,012***	***135,255,796***	-	-	-	-	-	-
EDTA	***135,252,061***	***135,253,582***	-	-	-	-	53,434,918	53,443,576
1700013H16Rik (Sycp3l)	-	-	-	-	-	-	53,742,901	53,757,970
Zfp36l3	-	-	-	-	-	-	53,772,686	53,776,394
Xlr	-	-	-	-	-	-	53,777,118	53,797,716
Gm14596	-	-	-	-	-	-	***54,349,070***	***54,365,937***
Gm14594	-	-	-	-	-	-	***54,440,097***	***54,457,011***
Gm16405 (duplicate of Slxl1)	-	-	-	-	-	-	***54,531,166***	***54,548,016***
Gm16430 (duplicate of Slxl1)	-	-	-	-	-	-	***54,622,169***	***54,639,019***
Gm14595	-	-	-	-	-	-	***54,713,174***	***54,730,013***
Gm14590	-	-	-	-	-	-	***54,810,820***	***54,827,792***

Table B3.3. Continued

Gene Name	Human Start	Human End	Cat Start	Cat End	Pig Start	Pig End	Mouse Start	Mouse End
Gm7950	-	-	-	-	-	-	***55,019,202***	***55,036,109***
Gm7958	-	-	-	-	-	-	***55,095,566***	***55,112,454***
Slxl1	-	-	-	-	-	-	***55,226,876***	***55,243,736***
Gm6660	-	-	-	-	-	-	***55,309,727***	***55,326,665***
Gm14625	-	-	-	-	-	-	***55,430,235***	***55,446,892***
Gm16404	-	-	-	-	-	-	***55,551,596***	***55,568,457***
Gm6664	-	-	-	-	-	-	55,798,252	55,816,984
3830403N18Rik (Xlrl)	-	-	-	-	-	-	55,839,252	56,153,496
Gm35586	-	-	-	-	-	-	55,891,713	55,895,852
Gm773	-	-	-	-	-	-	56,189,827	56,212,886
Gm35953	-	-	-	-	-	-	56,277,085	56,293,278
ZNF449	***135,344,098***	***135,363,413***	113,961,452	113,985,242	110,928,168	110,948,607	56,346,381	56,365,674
SMIM10L2A	***135,421,933***	***135,428,074***	114,013,978	114,014,619	110,959,731	110,963,079	56,374,586	56,378,470
INTS6L	***135,520,630***	***135,582,535***	114,101,290	114,157,328	111,011,860	111,072,939	56,454,825	56,507,843
CT45A1	***135,708,398***	***135,723,318***	-	-	-	-	-	-
CT45A3	***135,760,067***	***135,768,222***	-	-	-	-	-	-
CT45A5	***135,777,130***	***135,785,512***	-	-	-	-	-	-
CT45A6	***135,794,687***	***135,802,755***	-	-	-	-	-	-
CT45A2	***135,811,979***	***135,820,062***	-	-	-	-	-	-
CT45A7	***135,829,229***	***135,837,317***	-	-	-	-	-	-
CT45A8	***135,846,497***	***135,854,588***	-	-	-	-	-	-
CT45A9	***135,863,418***	***135,871,812***	-	-	-	-	-	-
CT45A10	***135,880,752***	***135,893,577***	-	-	-	-	-	-
SAGE1	135,893,607	135,913,061	-	-	-	-	-	-
LOC102901637 (INTS6LB)	-	-	114,182,695	114,188,831	-	-	-	-

Table B3.3. Continued

Gene Name	Human Start	Human End	Cat Start	Cat End	Pig Start	Pig End	Mouse Start	Mouse End
LOC102900259 (INTS6LC)	-	-	114,222,292	114,230,935	-	-	-	-
LOC102901675 (INTS6LD)	-	-	114,259,526	114,268,867	-	-	-	-
LOC100157667 (INTS6L)	-	-	-	-	111,094,016	111,110,838	-	-
Gm10477	-	-	-	-	-	-	56,524,742	56,525,432
MMGT1	135,962,072	135,973,975	114,297,265	114,308,362	111,145,712	111,158,298	56,585,512	56,597,919
SLC9A6	135,974,597	136,047,269	114,308,677	114,372,118	111,176,461	111,233,108	56,609,556	56,664,230
FHL1	136,146,702	136,211,359	114,449,186	114,511,042	111,309,989	111,367,900	56,731,720	56,793,346
MAP7D3	136,213,220	136,256,482	114,514,469	114,548,342	111,371,538	111,420,230	56,794,833	56,822,490
LOC100621767 (ADGRG4L)	-	-	-	-	111,469,540	111,488,383	-	-
LOC101083152	-	-	114,564,281	114,567,927	-	-	-	-
ADGRG4	136,300,963	136,416,888	114,586,469	114,674,561	111,490,544	111,534,000	56,894,234	56,980,270
BRS3	136,487,966	136,492,439	114,737,012	114,742,378	111,615,695	111,620,124	57,043,074	57,048,758
HTATSF1	136,496,932	136,512,346	114,746,098	114,760,739	111,624,303	111,643,566	57,053,570	57,067,183
VGLL1	136,532,152	136,556,807	114,776,946	114,798,448	111,662,565	111,684,676	57,088,106	57,106,540
CD40LG	136,648,177	136,660,390	114,884,203	114,894,916	111,778,666	111,788,901	57,212,143	57,224,042
ARHGEF6	136,665,547	136,781,656	114,906,203	115,006,306	111,796,485	111,917,195	57,231,485	57,338,758
RBMX	136,869,194	136,880,780	115,034,896	115,043,888	111,965,859	111,974,544	57,383,348	57,393,045
LOC101083916 (TM9SF2-like)	-	-	115,069,490	115,160,048	-	-	57,400,767	57,488,773
Pig Ampliconic	-	-	-	-	***111,987,560***	***112,007,873***	-	-
GPR101	137,030,148	137,031,674	115,162,846	115,170,113	112,109,592	112,111,333	57,496,668	57,503,757
ZIC3	137,566,187	137,577,691	115,600,122	115,615,061	112,592,951	112,605,087	58,021,464	58,041,735
LOC102158723 (Porcine ERV)	-	-	-	-	112,845,408	112,848,002	-	-
LOC110257694 (PRB2-like)	-	-	-	-	113,340,076	113,341,321	-	-
LOC101085187 (MAGEA13P)	-	-	116,243,719	116,248,096	113,340,314	113,343,205	58,911,461	58,920,304
FGF13	138,631,573	139,222,889	116,338,148	116,783,106	113,460,191	113,955,691	59,062,139	59,585,572

Table B3.3. Continued

Gene Name	Human Start	Human End	Cat Start	Cat End	Pig Start	Pig End	Mouse Start	Mouse End
LOC100523110 (RPS27L)	-	-	-	-	114,070,842	114,071,361	-	-
F9	139,530,720	139,563,459	117,080,317	117,111,828	114,218,704	114,250,436	59,999,464	60,030,760
MCF2	139,581,768	139,708,279	117,127,818	117,235,327	114,258,082	114,356,106	60,055,956	60,179,184
ATP11C	139,726,346	139,933,083	117,268,592	117,440,354	114,394,436	114,575,658	60,223,283	60,404,861
LOC110257749 (SEC31L)	-	-	-	-	114,575,980	115,026,005	-	-
CXorf66	139,955,725	139,965,518	117,463,792	117,466,922	114,596,809	114,602,521	60,436,052	60,456,195
LOC389895 (C16orf72)	140,091,667	140,092,911	117,587,366	117,588,627	114,686,410	114,687,790	60,548,008	60,549,217
LOC102899102 (TNP4L)	-	-	117,716,423	117,717,176	-	-	-	-
SOX3	140,502,987	140,505,060	117,894,272	117,896,500	115,017,704	115,020,601	60,891,366	60,893,430
CDR1	140,783,260	140,784,558	118,142,324	118,165,591	115,221,975	115,223,194	61,183,246	61,185,558
LOC101086956 (MAGE10L)	-	-	118,221,035	118,228,465	115,168,447	115,304,951	61,288,157	61,292,488
SPANXB1	***141,002,591***	***141,003,706***	-	-	-	-	-	-
LDOC1	***141,175,745***	***141,177,214***	118,564,662	118,566,992	115,580,622	115,583,129	61,709,616	61,710,950
SPANXC	***141,241,463***	***141,242,550***	-	-	-	-	-	-
SPANXA1	***141,583,674***	***141,584,738***	-	-	-	-	-	-
SPANXA2	141,589,708	141,590,772	-	-	-	-	-	-
LOC645188	141,625,864	141,626,733	-	-	-	-	-	-
SPANXD	141,697,401	141,698,739	-	-	-	-	-	-
MAGEC3	141,838,316	141,897,832	-	-	-	-	-	-
MAGEC1	141,903,856	141,909,401	-	-	-	-	-	-
MAGEC2	142,202,342	142,205,290	-	-	-	-	-	-
SPANXN4	143,025,918	143,038,637	-	-	-	-	-	-
SPANXN3	143,508,735	143,517,475	-	-	-	-	-	-
LOC110257823	-	-	-	-	115,725,317	115,728,246	-	-
4933402E13Rik	-	-	-	-	-	-	62,282,960	62,292,084

Table B3.3. Continued

Gene Name	Human Start	Human End	Cat Start	Cat End	Pig Start	Pig End	Mouse Start	Mouse End
Tslrn1	-	-	-	-	-	-	62,510,539	62,527,011
3830417A13Rik	-	-	-	-	-	-	64,173,548	64,178,778
SLITRK4	143,622,790	143,636,107	120,223,320	120,234,950	117,145,647	117,157,275	64,264,883	64,277,394
SPANXN2	143,712,035	143,720,668	-	-	-	-	-	-
SPANXN1	145,247,587	145,256,208	-	-	-	-	-	-
LOC102157658 (MAGEA13P)	-	-	-	-	118,448,490	118,566,819	-	-
LOC100511430 (TNP4)	-	-	-	-	118,621,838	118,622,598	-	-
SLITRK2	145,817,829	145,829,856	121,945,879	121,954,118	118,826,531	118,837,134	66,649,318	66,661,402
CXorf51B	***146,809,784***	***146,810,411***	-	-	-	-	-	-
CXorf51A	***146,814,104***	***146,814,731***	-	-	-	-	-	-
LOC105373347 (PPHLN1)	147,181,049	147,271,945	-	-	-	-	-	-
LOC111558724 (FH14)	-	-	122,380,408	122,383,513	-	-	-	-
LOC101091265 (ELL2)	-	-	122,482,843	122,491,048	-	-	-	-
LOC102901550 (TNP4)	-	-	122,650,566	122,651,363	-	-	-	-
Gm1140	-	-	-	-	-	-	***67,682,900***	***67,693,562***
Gm14692	-	-	-	-	-	-	67,695,698	67,706,258
4933436I01Rik	-	-	-	-	-	-	67,919,864	67,921,450
FMR1	147,911,951	147,951,127	123,587,612	123,633,978	120,361,232	120,400,282	68,678,541	68,717,963
FMR1NB	147,981,329	148,026,667	123,647,449	123,682,019	120,421,266	120,458,866	68,761,839	68,804,561
Gm6812	-	-	-	-	-	-	68,892,373	68,893,053
AFF2	148,500,619	149,000,663	124,050,517	124,511,205	120,777,559	121,263,574	69,360,331	69,872,054
LOC110257797	-	-	-	-	121,231,431	121,500,019	-	-
1700020N15Rik	-	-	-	-	-	-	69,945,281	69,945,980
IDS	149,476,990	149,505,354	124,824,846	124,865,058	121,572,797	121,596,525	70,343,069	70,385,816
IDS2	149,511,554	149,526,523	-	-	-	-	-	-

Table B3.3. Continued

Gene Name	Human Start	Human End	Cat Start	Cat End	Pig Start	Pig End	Mouse Start	Mouse End
CXorf40A	149,540,630	149,555,338	124,866,965	124,871,751	***121,620,738***	***121,628,536***	70,385,913	70,389,416
HSFX3	149,548,210	149,549,852	-	-	-	-	-	-
MAGEA9B	***149,581,653***	***149,587,468***	-	-	-	-	-	-
LOC111558690 (BRP-like)	-	-	124,871,180	124,888,322	-	-	-	-
LOC110257715 (MAGEA10L)	-	-	-	-	***121,632,759***	***121,634,665***	-	-
HSFX2	***149,592,512***	***149,595,310***	125,151,828	125,155,127	121,645,509	121,647,655	-	-
TMEM185A	***149,596,556***	***149,631,912***	125,113,727	125,150,354	121,648,807	121,682,484	***70,459,753***	***70,477,190***
MAGEA11	***149,688,202***	***149,717,268***	-	-	-	-	-	-
HSFX1	***149,774,068***	***149,776,859***	-	-	-	-	-	-
MAGEA9	***149,781,921***	***149,787,737***	-	-	-	-	-	-
MAGEA8	149,881,141	149,885,835	-	-	-	-	154,985,775	154,995,850
LOC100156285 (MAGEA10L)	-	-	-	-	121,688,350	121,691,552	-	-
LOC100514940 (MAGEA10L)	-	-	-	-	121,699,620	121,702,252	-	-
LOC100513838 (MAGEA10L)	-	-	-	-	121,708,847	121,709,729	-	-
LOC100513259 (MAGE8L)	-	-	-	-	121,755,015	121,756,449	-	-
LOC106504240 (MAGE10L)	-	-	-	-	121,763,484	121,765,954	-	-
CXorf40B	149,924,152	149,938,850	125,300,573	125,305,176	121,769,595	121,777,428	-	-
HSFX4	149,929,624	149,931,287	125,297,835	125,299,840	121,769,734	121,771,155	-	-
LOC106506163	-	-	-	-	121,776,859	121,800,481	-	-
MAMLD1	150,361,422	150,514,178	125,702,650	125,823,203	122,120,275	122,233,692	71,050,040	71,156,056
MTM1	150,562,658	150,673,143	125,870,997	125,968,998	122,286,904	122,379,302	71,210,767	71,315,413
MTMR1	150,692,997	150,765,103	125,985,454	126,049,544	122,394,768	122,455,266	71,364,760	71,419,196
CD99L2	150,766,336	150,898,816	126,052,675	126,148,903	122,456,668	122,535,989	71,420,060	71,492,869
HMGB3	150,980,508	150,990,775	126,205,577	126,214,962	122,591,413	122,599,496	71,555,917	71,560,673
GPR50	151,176,584	151,182,855	126,316,402	126,321,654	122,708,551	122,713,163	71,663,667	71,669,257

Table B3.3. Continued

Gene Name	Human Start	Human End	Cat Start	Cat End	Pig Start	Pig End	Mouse Start	Mouse End
VMA21	151,396,555	151,409,364	126,486,696	126,497,037	122,872,141	122,882,976	71,812,753	71,824,706
PASD1	151,563,535	151,676,739	126,588,372	126,705,892	-	-	-	-
PRRG3	151,694,606	151,701,591	126,725,563	126,738,644	123,085,511	123,095,589	71,962,157	71,972,722
FATE1	151,716,036	151,723,194	126,733,107	126,757,827	123,106,086	123,113,178	71,972,986	71,989,046
CNGA2	151,734,746	151,745,564	126,772,037	126,779,800	123,115,980	123,130,043	71,991,849	72,010,218
MAGEA4	151,912,889	151,925,170	-	-	-	-	72,221,015	72,222,964
LOC111558691 (MAGE10L)	-	-	126,982,310	126,984,839	-	-	-	-
LOC105260348 (MAGE9L)	-	-	126,985,807	126,987,140	-	-	-	-
GABRE	151,953,124	151,974,679	127,030,023	127,047,691	123,278,459	123,296,070	72,256,926	72,274,822
MAGEA10-MAGEA5	152,114,049	152,138,578	-	-	-	-	-	-
MAGEA5	152,114,049	152,117,939	-	-	-	-	155,053,061	155,063,151
MAGEA10	152,133,310	152,138,578	-	-	123,400,803	123,401,915	72,381,870	72,386,858
LOC111558784 (MAGE8L)	-	-	127,126,566	127,128,537	-	-	-	-
GABRA3	152,166,234	152,451,359	127,260,126	127,630,283	123,456,952	123,684,814	72,432,676	72,656,890
GABRQ	152,637,895	152,659,827	127,819,235	127,840,253	123,834,884	123,855,187	72,825,163	72,842,607
MAGEA3	***152,698,742***	***152,702,347***	-	-	-	-	154,948,180	154,959,031
CSAG2	***152,708,261***	***152,713,861***	-	-	-	-	-	-
MAGEA2B	***152,714,529***	***152,718,607***	-	-	-	-	-	-
CSAG1	***152,727,484***	***152,733,736***	-	-	-	-	-	-
MAGEA12	***152,733,779***	***152,737,669***	-	-	-	-	-	-
MAGEA2	***152,749,863***	***152,753,942***	-	-	-	-	155,027,201	155,033,313
CSAG3	***152,753,921***	***152,760,222***	-	-	-	-	-	-
MAGEA6	***152,766,136***	***152,769,729***	-	-	-	-	154,923,895	-
LOC111558732 (*MAGE9L*)	-	-	127,862,926	127,864,162	-	-	-	-
LOC110257698	-	-	-	-	123,869,573	123,870,559	-	-

Table B3.3. Continued

Gene Name	Human Start	Human End	Cat Start	Cat End	Pig Start	Pig End	Mouse Start	Mouse End
CETN2	***152,827,327***	***152,830,757***	127,897,954	127,902,366	123,902,240	123,906,046	72,913,565	72,918,344
NSDHL	***152,830,967***	***152,869,363***	127,902,522	127,932,116	123,906,130	123,921,935	72,918,521	72,958,528
ZNF185	***152,898,132***	***152,973,481***	127,932,370	128,027,995	123,942,981	123,988,102	72,987,247	73,031,543
PNMA5	***152,988,824***	***152,994,458***	-	-	123,998,935	124,001,201	73,033,981	73,040,260
PNMA6FL	-	-	128,050,943	128,059,895	-	-	-	-
PNMA3	***153,056,343***	***153,060,467***	128,097,870	128,102,139	-	-	73,064,787	73,068,191
PNMA6A	***153,072,414***	***153,075,019***	X_random	X_random	124,005,969	124,007,679	-	-
MAGEA1	***153,179,284***	***153,183,878***	-	-	-	-	155,088,519	155,089,790
Cat Ampliconic	-	-	***128,107,195***	***128,118,411***	-	-	-	-
LOC111558739 (SYCP3L)	-	-	128,189,405	128,197,106	***124,048,096***	***124,060,530***	-	-
LOC100154655 (SYCP3L)	-	-	-	-	124,078,500	124,089,002	-	-
Xlr4a	-	-	-	-	-	-	73,074,343	73,082,526
Xlr3a	-	-	-	-	-	-	73,086,293	73,097,095
Xlr5a	-	-	-	-	-	-	73,107,621	73,119,113
LOC111558692 (MUM1L1L)	-	-	128,234,075	128,234,782	124,109,383	124,111,245	***73,117,048***	***73,129,269***
LOC111558735 (MUM1L1L)	-	-	128,235,772	128,237,735	-	-	-	-
PNMA6F	153,317,681	153,321,822	128,303,040	128,309,091	124,144,682	124,147,565	-	-
DXBay18	-	-	-	-	-	-	***73,137,223***	***73,149,450***
Xlr5b	-	-	-	-	-	-	***73,147,383***	***73,158,886***
Spin2d	-	-	-	-	-	-	73,175,302	73,176,989
Xlr3b	-	-	-	-	-	-	73,192,150	73,202,930
Xlr4b	-	-	-	-	-	-	73,214,321	73,222,453
F8a	-	-	-	-	-	-	73,228,306	73,230,795
Xlr4c	-	-	-	-	-	-	73,234,076	73,243,130
Xlr3c	-	-	-	-	-	-	73,254,538	73,265,576

Table B3.3. Continued

Gene Name	Human Start	Human End	Cat Start	Cat End	Pig Start	Pig End	Mouse Start	Mouse End
Xlr5c	-	-	-	-	-	-	73,283,361	73,293,970
ZNF275	153,334,155	153,352,926	128,326,051	128,343,093	124,156,708	124,173,566	73,342,621	73,359,080
LOC110257703 (PNMA6FL)	-	-	-	-	124,203,040	124,204,023	-	-
ZFP92	153,411,461	153,426,481	128,407,059	128,412,830	124,215,906	124,227,853	73,410,978	73,428,382
TREX2	153,444,720	153,446,487	128,430,950	128,434,444	124,234,736	124,244,193	73,433,705	73,435,343
HAUS7	153,447,665	153,495,525	128,434,483	128,460,268	124,234,735	124,258,133	73,437,315	73,459,029
ECM2	153,483,920	153,487,088	128,469,239	128,472,361	124,266,197	124,274,910	-	-
BGN	153,494,889	153,509,554	128,484,343	128,498,998	124,282,048	124,295,450	73,483,601	73,495,936
ATP2B3	153,517,677	153,582,929	128,505,832	128,568,243	124,318,874	124,357,653	73,503,023	73,573,275
CCNQ	153,587,925	153,599,177	128,572,284	128,583,798	124,361,003	124,370,072	-	-
DUSP9	153,642,443	153,651,326	128,633,805	128,636,658	124,404,852	124,414,246	73,639,389	73,643,514
PNCK	153,669,723	153,687,568	128,653,120	128,661,442	124,425,867	124,430,000	73,655,992	73,660,111
SLC6A8	153,688,297	153,696,593	128,665,542	128,674,396	124,435,282	124,452,182	73,673,133	73,682,502
BCAP31	153,700,492	153,724,746	128,678,321	128,706,934	124,456,999	124,484,923	73,686,178	73,717,815
ABCD1	153,724,851	153,744,762	128,707,316	128,727,582	124,485,076	124,501,742	73,716,597	73,738,534
PLXNB3	153,764,196	153,779,346	128,727,845	128,755,120	124,507,953	124,522,780	73,756,559	73,772,510
SRPK3	153,781,001	153,785,732	128,756,632	128,761,273	124,524,102	124,528,607	73,773,994	73,778,925
IDH3G	153,785,766	153,794,523	128,761,290	128,772,140	124,528,643	124,537,570	73,778,963	73,787,008
SSR4	153,794,175	153,798,512	128,772,361	128,776,474	124,537,545	124,541,635	73,787,028	73,790,828
PDZD4	153,802,166	153,830,567	128,779,722	128,811,573	124,543,845	124,567,477	73,793,357	73,825,033
L1CAM	153,861,514	153,886,174	128,861,436	128,876,163	124,595,649	124,618,292	73,853,780	73,880,838
LCA10	153,880,672	153,888,990	-	-	-	-	-	-
AVPR2	153,902,531	153,907,166	128,900,470	128,905,638	124,638,433	124,640,969	73,891,798	73,894,428
ARHGAP4	153,907,376	153,926,260	128,905,868	128,921,967	124,641,222	124,655,312	73,894,352	73,911,346
NAA10	153,929,827	153,935,154	128,927,017	128,931,766	124,658,429	124,662,744	73,916,870	73,921,944

Table B3.3. Continued

Gene Name	Human Start	Human End	Cat Start	Cat End	Pig Start	Pig End	Mouse Start	Mouse End
RENBP	153,935,263	153,944,780	128,932,038	128,939,112	124,662,957	124,670,522	73,922,121	73,930,850
HCFC1	153,947,556	153,971,833	128,941,002	128,965,738	124,673,125	124,694,846	73,942,792	73,967,135
TMEM187	153,972,540	153,983,195	128,966,387	128,971,279	124,695,050	124,699,349	-	-
IRAK1	154,010,506	154,019,906	128,996,227	129,005,192	124,717,604	124,724,855	74,013,914	74,023,921
MECP2	154,021,800	154,097,731	129,007,461	129,075,745	124,735,523	124,789,063	74,026,592	74,085,690
OPN1LW	***154,144,224***	***154,159,032***	129,113,441	129,125,948	124,816,014	124,828,225	-	-
OPN1MW	***154,182,596***	***154,196,861***	-	-	-	-	74,127,466	74,150,756
OPN1MW2	***154,219,734***	***154,233,286***	-	-	-	-	-	-
OPN1MW3	***154,257,620***	***154,271,068***	-	-	-	-	-	-
TEX28	***154,271,265***	***154,295,356***	129,126,064	129,142,270	124,828,222	124,840,093	74,150,909	74,167,838
TKTL1	***154,295,674***	***154,330,363***	129,147,421	129,172,902	124,840,072	124,867,313	74,177,259	74,208,498
LOC110257704 (COL3A1)	-	-	-	-	124,870,296	124,881,007	-	-
Cat Ampliconic	-	-	***129,175,617***	***129,187,033***	-	-	-	-
FLNA	***154,348,529***	***154,374,638***	129,188,026	129,214,016	124,889,934	124,915,000	74,223,461	74,249,852
EMD	***154,379,237***	***154,381,523***	129,218,429	129,221,255	124,883,047	124,885,429	74,254,672	74,261,563
RPL10	***154,398,065***	***154,402,339***	129,238,112	129,240,764	124,927,249	124,929,791	74,270,816	74,273,135
DNASE1L1	***154,401,236***	***154,412,101***	129,241,643	129,247,538	124,931,410	124,937,569	74,273,217	74,282,333
TAZ	***154,411,406***	***154,421,726***	129,249,620	129,261,266	124,940,036	124,946,275	74,281,900	74,295,319
ATP6AP1	***154,428,632***	***154,436,517***	129,270,220	129,278,150	124,952,532	124,960,344	74,294,907	74,304,721
GDI1	***154,436,913***	***154,443,467***	129,278,763	129,284,794	124,960,835	124,966,833	74,305,012	74,311,867
FAM50A	***154,444,126***	***154,450,654***	129,285,560	129,292,455	124,967,821	124,973,487	74,313,033	74,320,149
PLXNA3	***154,458,281***	***154,473,646***	129,304,283	129,319,179	124,982,307	124,997,098	74,329,050	74,344,689
LAGE3	***154,477,769***	***154,479,257***	129,326,726	129,329,082	125,001,673	125,003,206	74,352,162	74,353,618
UBL4A	***154,483,717***	***154,486,670***	129,336,966	129,339,605	125,007,121	125,009,657	74,365,718	74,368,548
Gm44504	-	-	-	-	-	-	74,367,447	74,373,349

Table B3.3. Continued

Gene Name	Human Start	Human End	Cat Start	Cat End	Pig Start	Pig End	Mouse Start	Mouse End
SLC10A3	***154,487,306***	***154,490,690***	129,340,271	129,344,010	125,010,128	125,013,992	74,369,217	74,373,349
FAM3A	***154,506,159***	***154,516,249***	129,347,921	129,360,128	125,015,985	125,025,119	74,384,719	74,393,274
G6PD	***154,531,390***	***154,547,586***	129,360,588	129,374,276	125,029,147	125,041,040	74,409,486	74,429,161
IKBKG	***154,542,240***	***154,565,046***	129,369,288	129,398,062	125,035,575	125,056,445	74,393,291	74,456,792
CTAG1A	***154,585,154***	***154,586,816***	-	-	-	-	-	-
CTAG1B	***154,617,609***	***154,619,271***	-	-	-	-	-	-
CTAG2	***154,651,972***	***154,653,579***	-	-	-	-	65,047,644	65,049,017
LOC105259690 (SLC6A8-like)	-	-	129,432,061	129,435,675	-	-	-	-
LOC100737597 (IF-2-like)	-	-	-	-	125,072,091	125,097,974	-	-
LOC102158155 (CXorf49L)	-	-	-	-	125,109,070	125,112,343	-	-
LOC110257705 (OR10VL)	-	-	-	-	125,122,994	125,125,168	-	-
LOC102166335 (LAGE3-like)	-	-	-	-	125,150,593	125,152,014	-	-
Gm6880 (Ctag2 duplicate)	-	-	-	-	-	-	74,480,117	74,481,544
Olfr1325	-	-	-	-	-	-	74,594,328	74,595,275
Gm5640	-	-	-	-	-	-	***74,639,114***	***74,646,110***
Gm6890	-	-	-	-	-	-	***74,740,927***	***74,742,188***
Gm5936	-	-	-	-	-	-	***74,836,722***	***74,843,369***
FCAB-331E2-_Transcript	-	-	129,435,118*	129,505,645*	-	-	-	-
FCAB-331E3_Transcript2	-	-	129,435,118*	129,505,645*	-	-	-	-
FCAB-331E3_Transcript3	-	-	129,435,118*	129,505,645*	-	-	-	-
FCAB-331E2_Transcript4	-	-	129,435,118*	129,505,645*	-	-	-	-
FCAB-331E3_Transcript5	-	-	129,435,118*	129,505,645*	-	-	-	-
GAB3	***154,675,249***	***154,751,583***	129,438,187	129,516,836	125,158,148	125,209,080	74,988,545	75,084,905
DKC1	***154,762,742***	***154,777,689***	129,529,084	129,540,735	125,218,928	125,228,881	75,095,854	75,109,777
MPP1	***154,778,684***	***154,805,527***	129,541,699	129,568,634	125,230,412	125,252,384	75,109,733	75,131,016

Table B3.3. Continued

Gene Name	Human Start	Human End	Cat Start	Cat End	Pig Start	Pig End	Mouse Start	Mouse End
SMIM9	***154,823,348***	***154,834,662***	129,575,941	129,584,496	125,254,737	125,262,113	75,146,057	75,163,756
F8	***154,835,788***	***155,022,723***	129,589,215	129,720,258	125,265,330	125,350,853	75,170,344	75,383,525
H2AFB1	***154,885,042***	***154,885,558***	129,602,756	129,603,293	125,277,575	125,280,752	117,014,757	117,015,104
F8A1	***154,886,360***	***154,888,061***	129,603,829	129,605,084	125,280,978	125,282,771	-	-
FUNDC2	***155,026,789***	***155,056,916***	129,720,382	129,740,090	125,354,032	125,366,029	75,382,399	75,397,158
CMC4	***155,061,622***	***155,071,272***	129,745,125	129,754,466	125,372,941	125,383,272	75,404,846	75,416,584
MTCP1	***155,064,034***	***155,071,272***	129,749,102	129,754,465	125,372,941	125,377,910	75,410,442	75,416,584
BRCC3	***155,071,403***	***155,126,766***	129,755,040	129,818,159	125,383,392	125,439,082	75,416,623	75,455,702
LOC102900281 (PICALM-like)	-	-	129,830,163	129,879,234	125,452,118	125,481,762	-	-
VBP1	***155,197,007***	***155,239,841***	129,882,978	129,904,054	125,485,357	125,506,781	75,514,297	75,534,946
RAB39B	***155,258,234***	***155,264,589***	129,917,125	129,923,045	125,520,764	125,528,253	75,572,045	75,578,231
CLIC2	***155,276,207***	***155,334,681***	129,964,472	129,993,447	125,546,451	125,560,059	-	-
H2AFB2	***155,380,706***	***155,381,299***	130,003,795	130,004,439	125,560,361	125,568,560	116,681,178	116,681,525
F8A2	***155,382,115***	***155,383,230***	130,011,939	130,013,212	-	-	-	-
H2AFB3	***155,459,419***	***155,459,935***	130,013,760	130,015,009			120,312,582	120,313,226
LOC106507061 (EVA1C-like)	-	-	-	-	125,579,523	125,595,031	-	-
TMLHE	***155,489,007***	***155,612,961***	130,024,499	130,099,345	125,601,920	125,673,716	X_random	X_random
SPRY3	155,612,565	155,782,459	130,261,024	130,272,792	125,744,643	125,764,342	X_random	X_random
VAMP7	155,881,280	155,943,769	130,366,193	130,423,584	125,863,266	125,917,232	X_random	X_random
IL9R	155,997,581	156,013,017	-	-	-	-	-	-
GTPBP6AL	-	-	130,463,634	130,469,229	-	-	-	-
PLCXD1L	-	-	130,473,426	130,502,832	-	-	-	-

Table B3.4. Novel X-linked genes with ORFs >= 400bp found within the BAC clone sequences incorporated into FelCat9.0.

Locus	9.1 Start	9.1 End	Testes 1 FPKM	Testes 2 FPKM	Cerebellum FPKM	Heart FPKM	Kidney FPKM	Lung FPKM	BLASTP hit	E value
FCAB-31J21-174L20_Transcript 1	46,147,611	46,149,337	NA	17.4	NA	3.4	7.6	42.9	No significant hit	NA
FCAB-205P3-309N18...127P13_ Transcript1	49,065,362	49,066,904	3459.0	NA	417.7	103.6	269.3	361.1	No significant hit	NA
FCAB-89F20_Transcript1	113,260,518	113,371,029	391.3	164.8	2.6	21.8	15.0	9.7	No significant hit	NA
FCAB-331E2_Transcript1	128,973,412	128,976,005	1712.9	174.3	373.1	124.4	359.2	431.8	No significant hit	NA
FCAB-331E2_Transcript2	128,981,795	128,986,452	NA	NA	NA	2.6	NA	18.2	Select seq XP_026911154.1 LOW QUALITY PROTEIN: olfactory receptor 3A3-like [*Acinonyx jubatus*]	4.00E-111
FCAB-331E2_Transcript3	129,004,897	129,011,236	NA	135.7	18.8	3.9	8.2	27.4	DUF1725 domain-containing protein [*Neisseria polysaccharea*]	4.00E-53
FCAB-331E2_Transcript4	129,011,378	129,012,810	135.9	16.2	13.0	3.9	8.2	27.4	Select seq XP_025785020.1 cancer/testis antigen 1-like [*Puma concolor*]	3.00E-46
FCAB-331E2_Transcript5	129,021,032	129,022,179	982.3	NA	4.2	NA	5.1	14.3	No significant hit	NA

Table B3.5. Coordinates for blocks of synteny shared between the ancestral X chromosome orders of the human, cat, and pig X chromosomes with the rearranged X chromosome of the mouse. Ampliconic regions falling within 1 Mb of synteny breakpoints are provided for each species.

Gene Name	Human Start	Human End	Cat Start	Cat End	Pig Start	Pig End	Mouse Start	Mouse End
MID1	10,445,310	11,111,264	7,719,203	8,316,004	7,235,386	7,906,049	169,685,199	169,990,798
SAT1	23,783,158	23,786,223	20,204,990	20,208,126	19,908,037	19,910,553	155,213,126	155,216,449
APOO	23,833,348	23,907,940	20,247,999	20,304,786	19,953,147	20,017,784	94,367,011	94,417,093
PRRG1	37,349,275	37,457,295	32,686,315	32,996,245	33,064,499	33,232,556	78,449,610	78,583,912
LANCL3	37,571,569	37,684,463	33,177,349	33,273,100	33,426,731	33,543,506	9,199,973	9,268,085
SPACA5	48,004,336	48,009,736	42,412,167	42,417,462	42,476,719	42,478,783	21,068,488	21,077,959
SLC38A5	48,458,537	48,470,256	42,545,288	42,558,731	42,590,608	42,601,537	8,271,381	8,280,179
NUDT11	51,490,011	51,496,607	44,936,298	44,942,817	44,912,160	44,918,977	6,047,507	6,054,751
GPR173	53,049,324	53,080,615	46,084,470	46,108,219	45,883,738	45,908,197	152,343,598	152,368,966
ALAS2	55,009,055	55,031,064	48,091,113	48,117,734	47,871,519	47,896,041	150,547,417	150,570,622
ZXDB	57,591,836	57,597,477	49,966,684	49,972,073	49,787,449	49,790,835	94,724,569	94,730,191
LUZP4	115,289,717	115,307,556	96,803,935	96,842,171	94,631,490	94,647,712	148,857,694	148,924,139
AGTR2	116,170,705	116,174,972	97,342,561	97,346,822	95,267,709	95,272,237	21,484,549	21,488,833
RAB39B	155,258,234	155,264,589	129,917,125	129,923,045	125,520,764	125,528,253	75,572,045	75,578,231

Table B3.5. Continued

Gene Name	Human Ampliconic Start	Human Ampliconic End	Cat Ampliconic Start	Cat Ampliconic End	Pig Ampliconic Start	Pig Ampliconic End	Mouse Ampliconic Start	Mouse Ampliconic End
MID1	NA	NA	NA	NA	NA	NA	170,737,096	170,881,298
SAT1	NA	NA	NA	NA	NA	NA	154,000,736	154,066,465
APOO	NA	NA	NA	NA	NA	NA	94,706,195	94,816,711
PRRG1	NA	NA	NA	NA	NA	NA	NA	NA
LANCL3	NA	NA	NA	NA	NA	NA	8,480,304	8,499,226
SPACA5	48,366,973	48,428,451	NA	NA	NA	NA	NA	NA
SLC38A5	48,366,973	48,428,451	NA	NA	NA	NA	8,480,304	8,499,226
NUDT11	51,668,020	52,920,401	44,753,753	46,069,961	45,054,183	45,858,289	5,210,613	5,313,997
GPR173	51,668,020	52,920,401	44,753,753	46,069,961	45,054,183	45,858,289	152,409,657	152,438,240
ALAS2	55,453,508	55,480,140	49,525,430	50,521,076	48,022,185	48,041,616	147,824,892	150,099,691
ZXDB	NA	NA	49,525,430	50,521,076	49,881,516	49,894,192	94,706,195	94,816,711
LUZP4	115,843,208	115,886,139	NA	NA	95,013,653	95,139,430	147,824,892	150,099,691
AGTR2	115,843,208	115,886,139	NA	NA	95,013,653	95,139,430	NA	NA
RAB39B	154,149,491	155,504,549	129,175,617	129,187,033	NA	NA	NA	NA

TableB3.6a. Ampliconic regions of the human X chromosome using strict cutoff (>99% identity, > 10kb).

Strict Coordinates			
Chr	**Start**	**End**	**Length**
X	3,821,643	3,938,401	116,758
X	48,366,973	48,396,313	29,340
X	48,401,508	48,428,451	26,943
X	49,528,395	49,605,060	76,665
X	51,668,020	51,692,962	24,942
X	52,040,715	52,077,119	36,404
X	52,473,395	52,502,273	28,878
X	52,669,254	52,729,002	59,748
X	52,756,104	52,789,250	33,146
X	52,881,815	52,920,401	38,586
X	55,453,508	55,480,140	26,632
X	58,716,337	58,735,885	19,548
X	58,763,534	58,804,502	40,968
X	58,805,114	58,846,090	40,976
X	59,467,608	59,477,725	10,117
X	59,822,836	59,838,106	15,270
X	59,934,949	59,975,932	40,983
X	61,974,100	61,984,719	10,619
X	63,129,000	63,185,599	56,599
X	63,207,431	63,224,416	16,985
X	71,683,102	71,740,753	57,651
X	71,740,843	71,798,461	57,618
X	72,740,931	72,860,189	119,258
X	72,860,608	72,979,901	119,293
X	102,197,558	102,338,169	140,611
X	102,402,578	102,489,558	86,980
X	103,953,836	103,989,433	35,597
X	106,266,280	106,278,816	12,536
X	115,843,208	115,886,139	42,931
X	120,038,203	120,087,100	48,897
X	120,162,296	120,198,162	35,866

TableB3.6b. Ampliconic regions of the human X chromosome clustering regions occurring within 500kb.

500kb Inclusive Coordinates			
Chr	**Start**	**End**	**Length**
X	3,821,643	3,938,401	116,758
X	48,366,973	48,428,451	61,478
X	49,528,395	49,605,060	76,665
X	51,668,020	52,920,401	1,252,381
X	55,453,508	55,480,140	26,632
X	58,716,337	58,846,090	129,753
X	59,467,608	59,975,932	508,324
X	61,974,100	61,984,719	10,619
X	63,129,000	63,224,416	95,416
X	71,683,102	71,798,461	115,359
X	72,740,931	72,979,901	238,970
X	102,197,558	102,489,558	292,000
X	103,953,836	103,989,433	35,597
X	106,266,280	106,278,816	12,536
X	115,843,208	115,886,139	42,931
X	120,038,203	120,198,162	159,959
X	120,929,382	120,987,033	57,651
X	135,116,125	135,889,734	773,609
X	141,002,115	141,585,255	583,140
X	146,801,758	146,822,756	20,998
X	149,573,132	149,796,256	223,124
X	152,678,581	153,293,926	615,345
X	154,149,491	155,504,549	1,355,058
		Total Length	6,804,303

Table B3.6a. Continued.

Strict Coordinates			
Chr	**Start**	**End**	**Length**
X	120,929,382	120,987,033	57,651
X	135,116,125	135,157,508	41,383
X	135,216,626	135,256,412	39,786
X	135,748,585	135,759,399	10,814
X	135,759,524	135,889,734	130,210
X	141,002,115	141,114,704	112,589
X	141,493,097	141,585,255	92,158
X	146,801,758	146,822,756	20,998
X	149,573,132	149,602,259	29,127
X	149,652,865	149,681,126	28,261
X	149,768,806	149,796,256	27,450
X	152,678,581	152,728,807	50,226
X	152,773,139	152,789,894	16,755
X	153,106,017	153,149,747	43,730
X	153,252,384	153,293,926	41,542
X	154,149,491	154,239,883	90,392
X	154,240,302	154,294,509	54,207
X	154,555,883	154,591,326	35,443
X	155,336,693	155,370,939	34,246
X	155,469,353	155,504,549	35,196
		Total Length	2,499,509

TableB3.7a. Ampliconic regions of the mouse X chromosome using strict cutoff (>99% identity, > 10kb).

Strict Coordinates			
Chr	**Start**	**End**	**Length**
X	3,321,011	3,347,932	26,921
X	3,348,201	3,431,872	83,671
X	3,439,804	3,451,412	11,608
X	3,476,117	3,495,015	18,898
X	3,495,017	3,556,354	61,337
X	3,660,458	3,704,118	43,660
X	3,714,075	3,779,736	65,661
X	3,782,690	3,827,901	45,211
X	3,923,993	3,961,313	37,320
X	3,962,137	4,111,011	148,874
X	4,136,367	4,201,051	64,684
X	4,258,658	4,443,883	185,225
X	4,598,904	4,618,418	19,514
X	4,742,211	4,843,849	101,638
X	4,904,458	4,960,180	55,722
X	5,210,613	5,221,813	11,200
X	5,257,527	5,313,997	56,470
X	8,480,304	8,499,226	18,922
X	11,298,326	11,328,120	29,794
X	13,167,131	13,197,566	30,435
X	24,808,958	24,822,949	13,991
X	24,922,951	25,116,508	193,557
X	25,280,935	25,299,474	18,539
X	25,332,935	25,346,966	14,031
X	25,994,206	26,009,345	15,139
X	26,159,602	26,198,225	38,623
X	26,210,800	26,229,312	18,512
X	26,251,237	26,266,016	14,779
X	26,267,438	26,279,245	11,807
X	26,299,005	26,420,395	121,390
X	26,433,807	26,472,439	38,632
X	26,715,257	26,727,071	11,814
X	26,917,216	27,055,754	138,538
X	27,205,756	27,499,669	293,913
X	27,549,671	28,023,371	473,700

TableB3.7b. Ampliconic regions of the mouse X chromosome clustering regions occurring within 500kb.

500Kb Inclusive Coordinates			
Chr	**Start**	**End**	**Length**
X	3,321,011	5,313,997	1,992,986
X	8,480,304	8,499,226	18,922
X	11,298,326	11,328,120	29,794
X	13,167,131	13,197,566	30,435
X	24,808,958	25,346,966	538,008
X	25,994,206	35,451,485	9,457,279
X	37,251,522	37,402,518	150,996
X	54,272,650	55,791,158	1,518,508
X	67,680,296	67,694,529	14,233
X	70,430,244	70,509,140	78,896
X	73,118,084	73,148,456	30,372
X	74,642,652	74,840,177	197,525
X	91,447,322	91,965,791	518,469
X	94,706,195	94,816,711	110,516
X	103,003,108	103,016,033	12,925
X	123,113,661	126,051,901	2,938,240
X	135,107,360	135,564,906	457,546
X	147,824,892	150,099,691	2,274,799
X	152,409,657	152,438,240	28,583
X	154,000,736	154,066,465	65,729
X	170,737,096	170,881,298	144,202
		Total Length	20,608,963

TableB3.7a. Continued

Chr	Start	End	Length
X	28,030,950	28,330,237	299,287
X	28,380,239	29,212,943	832,704
X	29,362,945	29,569,352	206,407
X	29,681,555	29,972,627	291,072
X	29,974,783	30,369,758	394,975
X	30,370,140	30,439,560	69,420
X	30,549,523	31,006,869	457,346
X	31,104,766	31,130,036	25,270
X	31,255,911	31,417,070	161,159
X	31,567,072	31,604,033	36,961
X	31,630,673	31,706,480	75,807
X	31,919,957	31,977,723	57,766
X	31,977,962	32,031,802	53,840
X	32,075,984	32,140,829	64,845
X	32,149,111	32,289,726	140,615
X	32,427,544	32,517,555	90,011
X	32,520,760	32,571,402	50,642
X	32,601,893	32,618,366	16,473
X	32,624,954	32,707,210	82,256
X	32,881,725	32,906,987	25,262
X	32,942,347	33,035,727	93,380
X	33,188,949	33,209,419	20,470
X	33,214,673	33,279,684	65,011
X	33,281,834	33,352,905	71,071
X	33,417,706	33,453,431	35,725
X	33,512,129	33,554,196	42,067
X	33,560,919	33,681,286	120,367
X	33,689,474	33,841,329	151,855
X	33,846,720	33,920,714	73,994
X	33,931,484	33,941,630	10,146
X	33,945,829	33,967,540	21,711
X	34,129,362	34,247,008	117,646
X	34,254,513	34,308,085	53,572
X	34,471,278	34,519,668	48,390
X	34,768,448	35,451,485	683,037
X	37,251,522	37,277,030	25,508
X	37,377,032	37,402,518	25,486

TableB3.7a. Continued

Chr	Start	End	Length
X	54,272,650	54,754,891	482,241
X	54,761,275	54,828,053	66,778
X	54,928,055	55,112,715	184,660
X	55,163,804	55,288,392	124,588
X	55,715,294	55,791,158	75,864
X	67,680,296	67,694,529	14,233
X	70,430,244	70,455,471	25,227
X	70,484,614	70,509,140	24,526
X	73,118,084	73,131,370	13,286
X	73,136,827	73,148,456	11,629
X	74,642,652	74,678,992	36,340
X	74,793,890	74,840,177	46,287
X	91,447,322	91,481,925	34,603
X	91,584,063	91,599,464	15,401
X	91,628,213	91,641,318	13,105
X	91,933,822	91,965,791	31,969
X	94,706,195	94,727,546	21,351
X	94,796,717	94,816,711	19,994
X	103,003,108	103,016,033	12,925
X	123,113,661	123,149,059	35,398
X	123,158,461	123,191,222	32,761
X	123,262,709	123,318,496	55,787
X	123,321,385	123,335,116	13,731
X	123,342,121	123,368,667	26,546
X	123,379,371	123,475,649	96,278
X	123,489,028	123,561,025	71,997
X	123,572,516	123,607,933	35,417
X	123,613,039	123,645,896	32,857
X	123,713,444	123,769,354	55,910
X	123,779,319	123,814,419	35,100
X	123,820,858	123,917,134	96,276
X	123,931,926	123,993,720	61,794
X	124,000,401	124,042,137	41,736
X	124,268,139	124,672,743	404,604
X	124,822,745	124,865,918	43,173
X	124,868,342	125,018,865	150,523
X	125,089,738	125,108,437	18,699

TableB3.7a. Continued

Chr	Start	End	Length
X	125,291,139	125,446,051	154,912
X	125,546,053	125,779,746	233,693
X	125,829,748	126,030,876	201,128
X	126,031,191	126,051,901	20,710
X	135,107,360	135,191,594	84,234
X	135,359,456	135,564,906	205,450
X	147,824,892	147,857,590	32,698
X	147,863,248	147,898,894	35,646
X	147,901,748	147,940,501	38,753
X	147,992,631	148,165,713	173,082
X	148,216,144	148,248,845	32,701
X	148,253,696	148,604,431	350,735
X	148,608,614	148,652,329	43,715
X	148,673,599	148,722,263	48,664
X	148,731,613	148,804,652	73,039
X	148,983,047	149,006,871	23,824
X	149,006,876	149,021,828	14,952
X	149,037,151	149,431,596	394,445
X	149,443,362	149,547,074	103,712
X	149,547,111	149,578,505	31,394
X	149,687,734	149,702,528	14,794
X	149,954,753	150,003,383	48,630
X	150,026,498	150,099,691	73,193
X	152,409,657	152,438,240	28,583
X	154,000,736	154,066,465	65,729
X	170,737,096	170,758,505	21,409
X	170,859,849	170,881,298	21,449
		Total Length	12,760,152

TableB3.8a. Ampliconic regions of the domestic cat X chromosome using strict cutoff (>99% identity, > 10kb).

Strict Coordinates			
Chr	**Start**	**End**	**Length**
X	14,366,026	14,383,190	17,164
X	44,753,753	44,778,748	24,995
X	44,943,196	44,960,108	16,912
X	44,961,744	44,978,650	16,906
X	45,072,675	45,088,411	15,736
X	45,526,631	45,568,507	41,876
X	45,569,024	45,610,882	41,858
X	45,714,596	45,731,898	17,302
X	46,018,442	46,041,220	22,778
X	46,049,664	46,069,961	20,297
X	46,815,413	46,832,194	16,781
X	48,269,298	48,280,001	10,703
X	48,344,477	48,365,705	21,228
X	48,404,253	48,421,824	17,571
X	48,427,932	48,445,490	17,558
X	49,525,430	49,549,869	24,439
X	49,818,008	49,844,171	26,163
X	50,062,399	50,079,096	16,697
X	50,122,392	50,149,932	27,540
X	50,230,324	50,253,430	23,106
X	50,293,836	50,316,707	22,871
X	50,329,586	50,354,754	25,168
X	50,414,399	50,449,633	35,234
X	50,449,735	50,488,945	39,210
X	50,510,994	50,521,076	10,082
X	52,521,561	52,531,618	10,057
X	61,812,206	61,824,805	12,599
X	61,937,403	61,956,858	19,455
X	61,958,045	61,977,473	19,428
X	64,216,861	64,230,511	13,650
X	67,748,352	67,765,210	16,858
X	72,528,457	72,546,612	18,155
X	72,546,714	72,564,863	18,149
X	73,282,593	73,317,648	35,055
X	73,318,443	73,353,509	35,066

TableB3.8b. Ampliconic regions of the domestic cat X chromosome clustering regions occurring within 500kb.

500Kb Inclusive Coordinates			
Chr	**Start**	**End**	**Length**
X	14,366,026	14,383,190	17,164
X	44,753,753	46,069,961	1,316,208
X	46,815,413	46,832,194	16,781
X	48,269,298	48,445,490	176,192
X	49,525,430	50,521,076	995,646
X	52,521,561	52,531,618	10,057
X	61,812,206	61,977,473	165,267
X	64,216,861	64,230,511	13,650
X	67,748,352	67,765,210	16,858
X	72,528,457	72,564,863	36,406
X	73,282,593	73,353,509	70,916
X	84,809,503	85,235,500	425,997
X	85,758,896	85,813,462	54,566
X	88,576,818	88,840,446	263,628
X	128,107,195	128,118,411	11,216
X	129,175,617	129,187,033	11,416
		Total Length	3,601,968

TableB3.8a. Continued

Chr	Start	End	Length
X	84,809,503	84,825,882	16,379
X	84,837,807	84,882,041	44,234
X	85,061,565	85,089,883	28,318
X	85,188,717	85,206,817	18,100
X	85,210,364	85,235,500	25,136
X	85,758,896	85,783,835	24,939
X	85,787,673	85,813,462	25,789
X	88,576,818	88,594,437	17,619
X	88,777,699	88,804,308	26,609
X	88,819,466	88,840,446	20,980
X	128,107,195	128,118,411	11,216
X	129,175,617	129,187,033	11,416
		Total Length	1,039,382

TableB3.9a. Ampliconic regions of the pig X chromosome using strict cutoff (>99% identity, > 10kb).

Strict Coordinates			
Chr	**Start**	**End**	**Length**
X	12,227,750	12,238,846	11,096
X	12,303,224	12,314,340	11,116
X	31,653,173	31,670,398	17,225
X	31,679,505	31,696,715	17,210
X	31,836,787	31,869,295	32,508
X	45,054,183	45,070,236	16,053
X	45,096,918	45,112,395	15,477
X	45,393,677	45,414,844	21,167
X	45,427,959	45,444,395	16,436
X	45,756,021	45,788,267	32,246
X	45,827,246	45,858,289	31,043
X	47,435,127	47,451,222	16,095
X	47,459,321	47,474,791	15,470
X	48,022,185	48,041,616	19,431
X	49,181,699	49,206,013	24,314
X	49,881,516	49,894,192	12,676
X	51,972,371	51,984,675	12,304
X	52,058,438	52,070,728	12,290
X	52,611,043	52,622,268	11,225
X	57,640,099	57,657,993	17,894
X	58,656,069	58,671,763	15,694
X	58,683,704	58,708,032	24,328
X	58,746,053	58,760,558	14,505
X	58,818,402	58,832,442	14,040
X	61,474,449	61,484,687	10,238
X	65,080,691	65,096,885	16,194
X	65,097,807	65,113,905	16,098
X	65,125,655	65,139,900	14,245
X	65,179,121	65,191,728	12,607
X	68,586,834	68,597,679	10,845
X	80,370,632	80,383,729	13,097
X	84,779,597	84,791,976	12,379
X	84,794,613	84,806,994	12,381
X	95,013,653	95,048,044	34,391
X	95,090,516	95,139,430	48,914

TableB3.8b. Ampliconic regions of the pig X chromosome clustering regions occurring within 500kb.

500Kb Inclusive Coordinates			
Chr	**Start**	**End**	**Length**
X	12,227,750	12,314,340	86,590
X	31,653,173	31,869,295	216,122
X	45,054,183	45,858,289	804,106
X	47,435,127	47,474,791	39,664
X	48,022,185	48,041,616	19,431
X	49,181,699	49,206,013	24,314
X	49,881,516	49,894,192	12,676
X	51,972,371	52,070,728	98,357
X	52,611,043	52,622,268	11,225
X	57,640,099	57,657,993	17,894
X	58,656,069	58,832,442	176,373
X	61,474,449	61,484,687	10,238
X	65,080,691	65,191,728	111,037
X	68,586,834	68,597,679	10,845
X	80,370,632	80,383,729	13,097
X	84,779,597	84,806,994	27,397
X	95,013,653	95,139,430	125,777
X	111,987,560	112,007,873	20,313
X	121,615,135	121,648,641	33,506
X	124,023,929	124,049,815	25,886
		Total Length	1,884,848

TableB3.9a. Continued

Chr	Start	End	Length
X	111,987,560	112,007,873	20,313
X	121,615,135	121,648,641	33,506
X	124,023,929	124,049,815	25,886
		Total Length	712,937

Table B3.10. Locations of novel genes on the human X chromosome. Genes within ampliconic regions are italicized, underlined, and in bold.

Gene Name	Human Start	Human End	Status
LOC105373102	1,392,349	1,396,881	SC
VCX3A	6,533,618	6,535,118	MC
VCX	7,842,262	7,844,143	MC
VCX2	8,169,944	8,171,267	MC
VCX3B	8,464,830	8,466,510	MC
FAM9A	8,790,795	8,801,383	MC
FAM9B	9,024,232	9,197,176	MC
FAM9C	13,035,617	13,044,798	MC
LOC105373133	13,044,544	13,086,975	SC
ATXN3L	13,318,647	13,320,399	MC
ZNF645	22,272,913	22,274,461	MC
SUPT20HL2	24,310,862	24,313,315	MC
SUPT20HL1	24,362,761	24,365,424	MC
DCAF8L1	27,977,993	27,981,449	MC
MAGEB2 (duplicate of MAGEB4)	30,215,558	30,220,089	MC
MAGEB1(duplicate of MAGEB4)	30,243,731	30,252,038	MC
FAM147A	34,129,752	34,132,330	MC
FAM47B	34,942,796	34,944,917	MC
LOC105373154 (MIB-like)	36,902,398	36,959,104	MC
LOC107985687 (EIF2L)	40,101,736	40,106,486	MC
KRBOX4	46,447,292	46,473,669	MC
SPACA5B (SPACA5 dup)	48,130,657	48,132,613	MC
SSX3	48,346,428	48,356,753	MC
SSX4	***48,383,527***	***48,393,343***	Ampliconic
SSX4B	***48,402,082***	***48,411,910***	Ampliconic
GAGE12J	49,322,030	49,329,387	MC
GAGE13	49,331,603	49,338,952	MC
GAGE12B	49,341,183	49,529,921	MC
GAGE12C	***49,532,177***	***49,539,541***	Ampliconic
GAGE12D	***49,541,722***	***49,549,107***	Ampliconic
GAGE12E	***49,551,289***	***49,558,649***	Ampliconic
GAGE12F	***49,560,808***	***49,568,208***	Ampliconic
GAGE12G	***49,570,400***	***49,577,757***	Ampliconic
GAGE12H	***49,579,949***	***49,587,304***	Ampliconic
GAGE2A	***49,589,515***	***49,596,827***	Ampliconic
GAGE1	***49,598,992***	***49,608,536***	Ampliconic

Table B3.10. Continued

Gene Name	Human Start	Human End	Status
PAGE1	49,687,450	49,696,019	MC
CENPVL2	***51,681,212***	***51,682,831***	Ampliconic
CENPVL1	***51,710,512***	***51,712,131***	Ampliconic
MAGED4B	***52,061,827***	***52,069,272***	Ampliconic
XAGE2	***52,368,996***	***52,375,683***	Ampliconic
XAGE1B	***52,495,668***	***52,500,815***	Ampliconic
XAGE1E (XAGE1E)	***52,512,074***	***52,517,221***	Ampliconic
SSX7	***52,644,061***	***52,654,900***	Ampliconic
SSX2	***52,696,896***	***52,707,227***	Ampliconic
SSX2B	***52,751,204***	***52,761,536***	Ampliconic
SPANXN5	***52,796,146***	***52,797,348***	Ampliconic
XAGE5	***52,811,295***	***52,818,301***	Ampliconic
XAGE3	***52,862,528***	***52,868,140***	Ampliconic
FAM156B	***52,891,599***	***52,908,560***	Ampliconic
PAGE2	55,089,052	55,092,836	MC
FAM104B	55,143,102	55,161,195	MC
MTRNR2L10	55,181,391	55,182,511	MC
PAGE5	55,220,347	55,224,108	MC
PAGE3	55,257,833	55,264,916	MC
SPIN2A	57,134,259	57,138,688	MC
ZXDA	57,905,430	57,910,633	MC
LOC105373242	69,031,120	69,084,509	SC
RAB41	70,280,978	70,285,088	MC
LOC105373244	70,427,358	70,436,800	SC
CXorf49B	***71,763,290***	***71,767,201***	Ampliconic
FLJ44635	72,144,184	72,161,750	MC
FAM236B	***72,781,865***	***72,782,660***	Ampliconic
FAM236D	***72,807,425***	***72,808,819***	Ampliconic
PABPC1L2B (Hominin-sp. duplicate)	73,003,513	73,005,712	MC
PGAM4	77,967,961	77,969,638	MC
RPA4	96,883,908	96,885,467	MC
LOC105373297 (ERVFC1)	97,841,749	97,849,381	MC
NXF2B	***102,360,395***	***102,440,008***	Ampliconic
RAB40AL	102,937,272	102,938,300	MC
RAB40A	103,493,266	103,519,489	MC
H2BFWT	104,011,147	104,013,687	MC
LOC101928589	110,175,773	110,177,788	SC
RBMXL3	115,189,400	115,192,868	MC
LOC105373314	115,454,266	115,465,553	SC

Table B3.10. Continued

Gene Name	Human Start	Human End	Status
RHOXF2B	***120,072,264***	***120,077,742***	Ampliconic
RHOXF2	***120,158,534***	***120,164,039***	Ampliconic
CT47B1	120,872,554	120,875,925	MC
CT47A12	120,877,490	120,932,399	MC
CT47A11	***120,933,840***	***120,937,260***	Ampliconic
CT47A10	***120,938,701***	***120,942,121***	Ampliconic
CT47A9	***120,943,561***	***120,946,981***	Ampliconic
CT47A8	***120,948,422***	***120,951,842***	Ampliconic
CT47A7	***120,953,282***	***120,956,600***	Ampliconic
CT47A6	***120,958,165***	***120,961,588***	Ampliconic
CT47A5	***120,963,026***	***120,966,446***	Ampliconic
CT47A4	***120,967,886***	***120,971,306***	Ampliconic
CT47A3	***120,972,746***	***120,976,166***	Ampliconic
CT47A2	***120,977,606***	***120,981,026***	Ampliconic
GLUD2	121,047,608	121,049,942	MC
TFDP3	133,216,669	133,218,348	MC
SMIM10L2B	135,094,985	135,098,634	MC
LOC101928677 (ETDAL)	***135,252,012***	***135,255,796***	Ampliconic
CT45A1	***135,708,398***	***135,723,318***	Ampliconic
CT45A3	***135,760,067***	***135,768,222***	Ampliconic
CT45A5	***135,777,130***	***135,785,512***	Ampliconic
CT45A6	***135,794,687***	***135,802,755***	Ampliconic
CT45A2	***135,811,979***	***135,820,062***	Ampliconic
CT45A7	***135,829,229***	***135,837,317***	Ampliconic
CT45A8	***135,846,497***	***135,854,588***	Ampliconic
CT45A9	***135,863,418***	***135,871,812***	Ampliconic
CT45A10	***135,880,752***	***135,893,577***	Ampliconic
SAGE1	135,893,607	135,913,061	MC
SPANXB1	***141,002,591***	***141,003,706***	Ampliconic
SPANXC	***141,241,463***	***141,242,550***	Ampliconic
SPANXA1	***141,583,674***	***141,584,738***	Ampliconic
SPANXA2	141,589,708	141,590,772	MC
LOC645188	141,625,864	141,626,733	SC
SPANXD	141,697,401	141,698,739	MC
MAGEC3	141,838,316	141,897,832	MC
MAGEC1	141,903,856	141,909,401	MC
MAGEC2	142,202,342	142,205,290	MC
SPANXN4	143,025,918	143,038,637	MC
SPANXN3	143,508,735	143,517,475	MC

Table B3.10. Continued

Gene Name	Human Start	Human End	Status
SPANXN2	143,712,035	143,720,668	MC
SPANXN1	145,247,587	145,256,208	MC
CXorf51B	***146,809,784***	***146,810,411***	Ampliconic
CXorf51A	***146,814,104***	***146,814,731***	Ampliconic
LOC105373347 (PPHLN1)	147,181,049	147,271,945	MC
IDS2	149,511,554	149,526,523	MC
HSFX3	149,548,210	149,549,852	MC
MAGEA9B	***149,581,653***	***149,587,468***	Ampliconic
MAGEA11	***149,688,202***	***149,717,268***	Ampliconic
HSFX1	***149,774,068***	***149,776,859***	Ampliconic
MAGEA9	***149,781,921***	***149,787,737***	Ampliconic
MAGEA10-MAGEA5	152,114,049	152,138,578	MC
CSAG2	***152,708,261***	***152,713,861***	Ampliconic
MAGEA2B	***152,714,529***	***152,718,607***	Ampliconic
CSAG1	***152,727,484***	***152,733,736***	Ampliconic
MAGEA12	***152,733,779***	***152,737,669***	Ampliconic
CSAG3	***152,753,921***	***152,760,222***	Ampliconic
LCA10	153,880,672	153,888,990	MC
OPN1MW2	***154,219,734***	***154,233,286***	Ampliconic
OPN1MW3	***154,257,620***	***154,271,068***	Ampliconic
CTAG1A	***154,585,154***	***154,586,816***	Ampliconic
CTAG1B	***154,617,609***	***154,619,271***	Ampliconic
IL9R	155,997,581	156,013,017	SC

Table B3.11. Locations of novel genes on the mouse X chromosome. Genes within ampliconic regions are italicized, underlined, and in bold.

Gene Name	Mouse Start	Mouse End	Status
Btbd35f23	3,076,920	3,078,823	MC
Btbd35f11	***3,441,731***	***3,443,690***	Ampliconic
Btbd35f18	***3,700,233***	***3,702,192***	Ampliconic
Btbd35f10	***3,750,931***	***3,752,885***	Ampliconic
Btbd35f16	***3,955,043***	***3,957,002***	Ampliconic
Btbd35f3	***4,037,674***	***4,039,622***	Ampliconic
Btbd35f28	***4,196,576***	***4,198,535***	Ampliconic
Btbd35f20	***4,289,286***	***4,291,245***	Ampliconic
Btbd35f17	***4,370,636***	***4,372,595***	Ampliconic
Btbd35f4	***4,800,411***	***4,802,359***	Ampliconic
Btbd35f27	***4,952,135***	***4,954,077***	Ampliconic
Mycs	5,466,904	5,469,265	SC
Btbd35f7	5,669,067	5,671,026	MC
Ssxb10	8,327,424	8,336,235	MC
Ssxb9	8,366,978	8,375,388	MC
Ssxb1	8,413,305	8,422,158	MC
Ssxb2	8,454,343	8,461,726	MC
Gm14459	8,513,543	8,526,726	MC
Ssxb6	8,542,604	8,548,255	MC
Ssxb3	8,583,404	8,591,810	MC
Ssxb7	8,639,607	8,645,722	MC
Ssxb8	8,684,843	8,690,891	MC
Ssx9	8,743,792	8,756,633	MC
Gm6592	8,843,994	8,849,607	MC
Gm5751	8,871,568	8,875,407	MC
Fthl17b	8,962,134	8,962,975	MC
Fthl17c	8,975,718	8,976,559	MC
Fthl17d	8,985,940	8,986,672	MC
Fthl17e	9,033,486	9,034,333	MC
Fthl17f	9,063,090	9,063,822	MC
4930402K13Rik (Fam47c duplicate)	9,104,562	9,106,342	MC
H2al1m	9,283,764	9,284,288	MC
H2al1k	9,350,599	9,351,137	MC
H2al1o	9,571,960	9,572,391	MC
H2al3	9,849,703	9,850,252	MC
H2al1a	***11,299,257***	***11,299,757***	Ampliconic

Table B3.11. Continued

Gene Name	Mouse Start	Mouse End	Status
H2a1b	***11,302,432***	***11,302,921***	Ampliconic
H2a1c	***11,305,655***	***11,305,972***	Ampliconic
H2a1d	***11,308,824***	***11,309,141***	Ampliconic
H2al1e	***11,311,934***	***11,312,427***	Ampliconic
H2a1f	***11,315,158***	***11,315,475***	Ampliconic
H2a1g	***11,318,256***	***11,318,764***	Ampliconic
H2a1h	***11,321,453***	***11,321,894***	Ampliconic
H2a1i	***11,324,659***	***11,324,976***	Ampliconic
H2a1j	***11,327,822***	***11,328,151***	Ampliconic
H2al1n	14,211,148	14,211,657	MC
Cypt1	16,522,869	16,523,691	MC
Tex13c3	19,479,424	19,481,319	MC
Gm39500	20,775,461	20,783,201	SC
Zfp300	21,079,150	21,089,700	MC
Ssxa1 (Ssxb paralog)	21,118,414	21,122,088	MC
Gm5932	21,169,950	21,179,036	SC
4930453H23Rik	21,206,580	21,221,672	SC
Gm4907	23,882,524	23,907,602	SC
Gm5934	24,474,308	24,499,163	MC
Gm4297	24,552,250	24,573,305	MC
Gm5935	24,753,162	24,775,164	MC
Gm5169	***25,277,487***	***25,301,455***	Ampliconic
Gm1993	25,548,501	25,570,449	MC
Gm5168	***26,028,214***	***26,051,041***	Ampliconic
Gm2012	***26,200,882***	***26,231,323***	Ampliconic
Gm2030	***26,287,199***	***26,310,044***	Ampliconic
Gm2036	***26,414,963***	***26,415,310***	Ampliconic
Slx	***26,522,657***	***26,545,565***	Ampliconic
Gm14525	***26,672,780***	***26,702,638***	Ampliconic
Gm6121	***26,912,743***	***26,935,669***	Ampliconic
Gm10230	***27,472,033***	***27,494,939***	Ampliconic
Gm10058	***27,924,021***	***27,946,885***	Ampliconic
Gm4836	***28,270,710***	***28,293,632***	Ampliconic
Gm10147	***28,584,995***	***28,607,874***	Ampliconic
Gm10096	***28,931,807***	***28,954,702***	Ampliconic
Gm10488	***29,976,314***	***29,999,162***	Ampliconic
Gm14632	***30,329,740***	***30,352,646***	Ampliconic
Gm10487	***30,820,542***	***30,843,469***	Ampliconic

Table B3.11. Continued

Gene Name	Mouse Start	Mouse End	Status
Btbd35f29	***31,117,576***	***31,119,415***	Ampliconic
LOC108168466 (SPIN2A-like)	***31,254,636***	***31,255,381***	Ampliconic
Btbd35f26	***31,336,219***	***31,338,442***	Ampliconic
Btbd35f19	***31,381,876***	***31,384,047***	Ampliconic
Spin2e	***31,618,191***	***31,619,833***	Ampliconic
Btbd35f25	***31,715,299***	***31,717,235***	Ampliconic
Gm21637	***31,960,926***	***31,962,821***	Ampliconic
Btbd35f22	***32,047,582***	***32,049,551***	Ampliconic
Gm5923	***32,158,701***	***32,160,371***	Ampliconic
Btbd35f2	***32,278,394***	***32,280,342***	Ampliconic
Gm21466	***32,437,174***	***32,438,844***	Ampliconic
Btbd35f1	***32,560,055***	***32,562,009***	Ampliconic
Gm5926	***32,649,778***	***32,651,426***	Ampliconic
Btbd35f21	***32,726,137***	***32,728,057***	Ampliconic
Btbd35f30	***32,894,527***	***32,896,366***	Ampliconic
Btbd35f5	***32,973,897***	***32,975,846***	Ampliconic
Btbd35f12	***33,313,338***	***33,315,288***	Ampliconic
Gm2854	***33,398,517***	***33,400,130***	Ampliconic
LOC100862179 (SPIN2A-like)	***33,476,982***	***33,478,575***	Ampliconic
Btbd35f15	***33,575,387***	***33,577,346***	Ampliconic
Btbd35f13	***33,657,140***	***33,659,082***	Ampliconic
Btbd35f14	***33,700,649***	***33,702,349***	Ampliconic
Gm46683	***33,777,828***	***33,778,538***	Ampliconic
Btbd35f6	***33,954,615***	***33,956,624***	Ampliconic
Btbd35f9	***34,159,642***	***34,161,734***	Ampliconic
LOC108168468 (SPIN2A-like)	***34,237,100***	***34,238,592***	Ampliconic
Btbd35f8	***34,450,295***	***34,452,516***	Ampliconic
Gm10486	***35,067,384***	***35,090,316***	Ampliconic
Gm14819	***35,404,215***	***35,427,094***	Ampliconic
Gm33383 (Ct47b-like)	36,401,263	36,403,446	MC
Gm6268 (Ct47-like)	36,403,477	36,404,654	MC
Gm9	37,208,502	37,211,041	SC
Rhox1	37,213,804	37,222,258	MC
Rhox2a	37,244,992	37,249,690	MC
Rhox3a	***37,249,919***	***37,258,978***	Ampliconic
Rhox4a	***37,265,067***	***37,270,136***	Ampliconic
Rhox3a2	***37,380,383***	***37,384,431***	Ampliconic
Rhox4a2	***37,390,860***	***37,395,621***	Ampliconic
Rhox2b	37,411,736	37,416,844	MC

Table B3.11. Continued

Gene Name	Mouse Start	Mouse End	Status
Rhox4b	37,432,494	37,437,278	MC
Rhox2c	37,453,480	37,458,417	MC
Rhox3c	37,469,868	37,473,961	MC
Rhox4c	37,480,337	37,485,124	MC
Rhox2d	37,493,410	37,497,715	MC
Rhox4d	37,514,385	37,519,176	MC
Rhox2e	37,530,322	37,541,851	MC
Rhox3e	37,546,975	37,550,897	MC
Rhox4e	37,557,412	37,562,283	MC
Rhox2f	37,571,421	37,576,159	MC
Rhox3f	37,581,352	37,585,496	MC
Rhox4f	37,602,895	37,607,686	MC
Rhox3g	37,623,431	37,628,139	MC
Rhox2g	37,639,112	37,643,470	MC
Rhox4g	37,646,501	37,651,327	MC
Rhox3h	37,657,688	37,668,764	MC
Rhox2h	37,668,997	37,673,277	MC
Rhox5	37,754,608	37,808,878	MC
Rhox6	37,827,055	37,829,857	MC
Rhox7a	37,831,686	37,841,171	MC
Rhox8	37,874,776	37,890,766	MC
Rhox9	37,899,097	37,901,770	MC
LOC102637882	37,940,614	37,945,671	MC
Rhox10	38,066,475	38,071,688	MC
Rhox11	38,076,598	38,085,139	MC
Rhox12	38,104,062	38,110,907	MC
Rhox13	38,120,840	38,129,966	MC
6030498E09Rik	38,772,780	38,962,688	SC
Cypt15	39,346,266	39,346,963	MC
Cypt14	39,862,919	39,863,604	MC
Olfr1320	49,683,504	49,684,463	MC
Olfr1321	49,726,973	49,727,932	MC
1700080O16Rik (Magea10L)	51,968,693	51,972,864	MC
1700013H16Rik (Sycp3l)	53,742,901	53,757,970	MC
Zfp36l3	53,772,686	53,776,394	MC
Xlr	53,777,118	53,797,716	MC
Gm14596	***54,349,070***	***54,365,937***	Ampliconic
Gm14594	***54,440,097***	***54,457,011***	Ampliconic
Gm16405 (duplicate of Slxl1)	***54,531,166***	***54,548,016***	Ampliconic

Table B3.11. Continued

Gene Name	Mouse Start	Mouse End	Status
Gm16430 (duplicate of Slxl1)	***54,622,169***	***54,639,019***	Ampliconic
Gm14595	***54,713,174***	***54,730,013***	Ampliconic
Gm14590	***54,810,820***	***54,827,792***	Ampliconic
Gm7950	***55,019,202***	***55,036,109***	Ampliconic
Gm7958	***55,095,566***	***55,112,454***	Ampliconic
Slxl1	***55,226,876***	***55,243,736***	Ampliconic
Gm6660	***55,309,727***	***55,326,665***	Ampliconic
Gm14625	***55,430,235***	***55,446,892***	Ampliconic
Gm16404	***55,551,596***	***55,568,457***	Ampliconic
Gm6664	55,798,252	55,816,984	MC
3830403N18Rik (Xlrl)	55,839,252	56,153,496	MC
Gm35586	55,891,713	55,895,852	MC
Gm773	56,189,827	56,212,886	SC
Gm35953	56,277,085	56,293,278	SC
Gm10477	56,524,742	56,525,432	SC
4933402E13Rik	62,282,960	62,292,084	MC
Tslrn1	62,510,539	62,527,011	MC
3830417A13Rik	64,173,548	64,178,778	MC
Gm1140	***67,682,900***	***67,693,562***	Ampliconic
Gm14692	67,695,698	67,706,258	MC
4933436I01Rik	67,919,864	67,921,450	SC
Gm6812	68,892,373	68,893,053	SC
1700020N15Rik	69,945,281	69,945,980	SC
Xlr4a	73,074,343	73,082,526	MC
Xlr3a	73,086,293	73,097,095	MC
Xlr5a	***73,107,621***	***73,119,113***	Ampliconic
DXBay18	***73,137,223***	***73,149,450***	Ampliconic
Xlr5b	***73,147,383***	***73,158,886***	Ampliconic
Spin2d	73,175,302	73,176,989	MC
Xlr3b	73,192,150	73,202,930	MC
Xlr4b	73,214,321	73,222,453	MC
F8a	73,228,306	73,230,795	MC
Xlr4c	73,234,076	73,243,130	MC
Xlr3c	73,254,538	73,265,576	MC
Xlr5c	73,283,361	73,293,970	MC
Gm44504	74,367,447	74,373,349	SC
Gm6880 (Ctag2 duplicate)	74,480,117	74,481,544	MC
Olfr1325	74,594,328	74,595,275	MC
Gm5640	***74,639,114***	***74,646,110***	Ampliconic

Table B3.11. Continued

Gene Name	Mouse Start	Mouse End	Status
Gm6890	***74,740,927***	***74,742,188***	Ampliconic
Gm5936	***74,836,722***	***74,843,369***	Ampliconic
Cldn34b3	76,264,306	76,267,673	MC
Cldn34b4	76,385,323	76,397,979	MC
Cldn34d	76,582,610	76,602,924	MC
Gm14744 (OBP duplicate)	77,864,758	77,870,033	MC
5430402E10Rik (Obp1-like)	77,919,786	77,925,062	MC
Obp1a	78,085,505	78,091,374	MC
Gm5938	78,125,467	78,130,403	MC
Obp1b	78,186,808	78,192,242	MC
Gm14743	78,240,462	78,245,921	MC
4930480E11Rik	78,369,643	78,371,128	SC
Gm8787	79,330,513	79,358,068	SC
Tsga8	83,486,674	83,487,924	MC
Samt3	86,044,199	86,047,313	MC
4930415L06Rik (PPP4RBL3like)	89,930,097	89,932,852	MC
Gm5072	91,392,500	91,393,534	MC
Gm14781	***91,624,249***	***91,635,671***	Ampliconic
LOC108168446 (MAGEB2L)	***91,625,353***	***91,626,477***	Ampliconic
Mageb1	92,015,191	92,016,333	MC
Gm5941	92,489,949	92,490,681	MC
1700010D01Rik	95,732,590	95,733,265	MC
Hsf3	96,306,178	96,456,294	MC
Gpr165	96,713,468	96,719,388	MC
Pgr15l	97,072,596	97,082,104	MC
Gm3858	101,818,196	101,821,816	SC
Gm9112	102,706,888	102,707,670	SC
LOC102632256 (Dmrtc1-like)	102,764,035	102,779,496	MC
Dmrtc1c1	102,797,880	102,809,950	MC
Dmrtc1c2	102,847,631	102,857,207	MC
1700031F05Rik (Dmrtc1-like)	102,859,166	102,866,353	MC
LOC102632367 (Dmrtc1-like)	102,880,507	102,882,317	MC
4930519F16Rik	103,182,991	103,256,632	SC
Cypt2	105,499,772	105,500,639	MC
LOC108168478 (RPL7A-like)	105,844,465	105,845,490	MC
Gm5127	106,583,151	106,710,557	MC
2610002M06Rik	107,782,751	107,816,334	SC
Gm732	107,945,735	107,948,436	SC
Gm6377	109,196,756	109,200,445	SC

Table B3.11. Continued

Gene Name	Mouse Start	Mouse End	Status
Gm34521	109,445,438	109,558,772	SC
4933403O08Rik	112,171,438	112,249,350	SC
Ube2dnl1	114,905,012	114,905,941	MC
Ube2dnl2	114,907,582	114,908,510	MC
Tgif2lx2	118,427,227	118,428,258	MC
Cldn34c1	***123,103,523***	***123,145,654***	Ampliconic
Astx6	***123,210,420***	***123,213,303***	Ampliconic
Srsx	***123,220,985***	***123,222,475***	Ampliconic
Gm2411	***123,346,785***	***123,378,718***	Ampliconic
Cldn34c2	***123,576,307***	***123,605,024***	Ampliconic
Cldn34c3	***123,788,212***	***123,823,149***	Ampliconic
Gm5945	***124,016,018***	***124,044,356***	Ampliconic
Vmn2r121	***124,127,339***	***124,135,910***	Ampliconic
4932411N23Rik	126,812,462	126,834,004	SC
Gm382	127,039,972	127,063,986	MC
Cldn34c4	127,721,175	127,736,554	MC
Gm15023	***135,167,470***	***135,177,653***	Ampliconic
Pramel3	***135,302,289***	***135,312,636***	Ampliconic
Gm5128	***135,373,250***	***135,383,393***	Ampliconic
Gm7903	***135,433,802***	***135,443,997***	Ampliconic
AV320801	***135,494,554***	***135,504,748***	Ampliconic
Nxf7	135,579,555	135,598,777	MC
Arxes1	136,033,367	136,034,946	MC
Kir3dl2	136,447,841	136,544,052	MC
Kir3dl1	136,517,999	136,534,306	MC
Tmsb15b2	136,954,988	136,957,979	MC
Tmsb15b1	136,974,022	136,976,874	MC
4930513O06Rik	139,086,243	139,093,403	SC
4933428M09Rik	139,179,215	139,181,009	SC
Trap1a	139,333,683	139,338,169	SC
Eif2c5	140,776,288	140,796,880	SC
GM15128	***147,789,653***	***147,833,116***	Ampliconic
GM15080	***147,992,993***	***148,033,174***	Ampliconic
GM15107	***148,180,675***	***148,224,710***	Ampliconic
GM15108	***148,488,949***	***148,521,446***	Ampliconic
GM15127	***148,663,588***	***148,703,849***	Ampliconic
Gm15093 (Ott-like duplicate)	***149,272,377***	***149,305,003***	Ampliconic
Gm15085 (Ott-like duplicate)	***149,447,617***	***149,487,784***	Ampliconic
Gm10439	***149,594,601***	***149,636,621***	Ampliconic

Table B3.11. Continued

Gene Name	Mouse Start	Mouse End	Status
Gm15097 (Ott-like duplicate)	***149,784,424***	***149,826,694***	Ampliconic
Gm15091 (Luzp4-like duplicate)	***149,941,387***	***149,984,982***	Ampliconic
Gm20153	150,157,426	150,177,319	SC
Gm5946 (Bmi1-like)	150,220,899	150,223,351	MC
Gm15104 (FAM156A/B-like)	150,311,924	150,336,719	MC
Gm15266	152,110,432	152,111,413	SC
Cypt3	153,558,593	153,559,435	MC
Kctd12b	153,685,154	153,696,386	MC
Samt1	***154,001,590***	***154,012,640***	Ampliconic
4921511M17Rik	***154,022,780***	***154,033,827***	Ampliconic
Gm10057	***154,043,986***	***154,055,041***	Ampliconic
Gm15143	***154,065,195***	***154,069,350***	Ampliconic
Gm15140	154,109,633	154,120,685	SC
4930524N10Rik	154,339,225	154,343,392	SC
Samt4	154,482,002	154,484,682	MC
Samt2	154,575,228	154,579,330	MC
Gm15142	154,635,586	154,639,130	SC
Cldn34b1	154,886,803	154,913,643	MC
Cldn34b2	155,124,917	155,147,849	MC
LOC108168453	161,297,856	161,334,722	SC
Rnf138rt1	163,760,139	163,761,332	MC
Siah1b	164,070,703	164,076,493	MC
Gm8817	167,063,449	167,070,370	SC
Gm6744	167,445,442	167,446,609	SC

Table B3.12. Locations of novel genes on the cat X chromosome. Genes within ampliconic regions are italicized, underlined, and in bold.

Gene Name	Cat Start	Cat End	Status
LOC101098289	85,641	98,815	SC
LOC102901245	134,637	198,935	SC
LOC111558781	199,003	207,366	SC
LOC102901718	1,000,391	1,023,639	SC
LOC109496913	1,135,898	1,139,304	SC
LOC111558770	1,156,778	1,159,735	SC
LOC109496425	7,099,539	7,124,318	SC
LOC101087815	7,315,456	7,316,115	SC
LOC101086788	13,225,739	13,226,878	SC
LOC105261399	13,277,526	13,280,230	SC
LOC102899587	15,116,750	15,122,803	SC
LOC101089581	15,755,581	15,774,652	SC
LOC101097703	20,032,463	20,032,949	SC
LOC101092829	20,037,757	20,038,541	SC
FELCATV1R3	21,603,090	21,603,992	MC
LOC105261400 (MAGEB4L)	22,294,290	22,295,724	MC
LOC111558670 (MAGEB17L)	22,330,125	22,331,507	MC
LOC493964 (MAGE)	23,749,406	23,751,663	MC
LOC109496467 (MAGEB2L)	26,047,369	26,054,530	MC
LOC109496468	26,204,512	26,480,297	SC
LOC111558783	33,312,823	33,313,387	SC
LOC101083747 (FTL-like)	34,224,027	34,292,046	MC
LOC101100947	43,153,840	43,155,056	SC
LOC111558700 (Formin-like protein 4)	***44,771,130***	***44,776,751***	Ampliconic
LOC109496513 (FMNL4L)	***44,942,817***	***44,947,984***	Ampliconic
LOC101088142	***45,159,674***	***45,160,706***	Ampliconic
LOC111558733 (MAGED4-like)	***45,708,172***	***45,712,838***	Ampliconic
LOC102899730 (PAGE3L)	***45,961,985***	***45,967,691***	Ampliconic
FCAB-31J21-174L20_Transcript1	***46,642,669****	***46,837,270****	Ampliconic
FCAB-205P3-309N18...127P13_Transcript1	***49,467,101***	***49,468,645***	Ampliconic
LOC105260211 (YBX3-like)	***50,479,533***	***50,487,761***	Ampliconic
LOC109496915	54,378,983	54,386,351	SC
LOC109496527 (RPS27-like)	55,015,556	55,015,836	MC
LOC101080617 (RPL37a-like)	55,302,143	55,302,411	MC
LOC102899556 (GCSH-like)	57,802,764	57,808,255	MC
LOC109496376	58,287,513	58,292,198	SC
DDX46	61,725,321	61,728,798	MC
LOC111558674 (FOXD1-like)	62,087,568	62,090,480	MC

Table B3.12. Continued

Gene Name	Cat Start	Cat End	Status
LOC111558772 (PNN-like)	63,580,928	63,584,015	MC
ERVK-6-like	65,966,702	65,974,312	MC
SH3BGRL3	68,986,150	68,989,752	MC
LOC109496330	70,625,536	70,683,139	SC
LOC102901599	***72,546,723***	***72,623,810***	Ampliconic
LOC102899136	72,624,536	72,629,477	S
LOC111558675	73,612,833	73,617,758	S
TEN1L	77,073,467	77,083,883	MC
LOC111558778	79,012,995	79,019,486	SC
LOC111558676	82,027,509	82,032,697	SC
ARMCX6L	83,751,767	83,763,976	MC
LOC111556301 (NFX2-like)	84,512,476	84,514,896	MC
LOC109496557	84,736,764	84,738,082	SC
LOC109496734 (ENVK7)	***84,973,344***	***84,977,226***	Ampliconic
LOC101101037 (SUMO2)	88,934,371	89,096,038	MC
LOC101085786 (RPS27-like)	89,933,289	89,933,540	MC
LOC111558744 (WH1-like)	89,988,732	89,989,977	MC
LOC102902491	92,055,713	92,058,745	SC
PERK10	96,453,223	96,454,335	MC
LOC111556295	96,454,338	96,459,055	SC
LOC102900824 (SET-like)	96,673,478	96,674,889	MC
LOC111558681	99,739,151	99,740,665	SC
LOC105259672 (CT47A-like)	99,740,979	99,743,939	MC
LOC111558775 (Rhox2b-like)	100,589,354	100,598,260	MC
LOC101093839 (RHOXF1)	100,643,391	100,676,554	MC
LOC111558717 (SLC25A5-like)	105,479,849	105,482,968	MC
LOC101098936 (COA1 homolog)	106,332,128	106,332,481	MC
LOC109496610 (COXB1-like)	112,978,811	112,979,341	MC
LOC102900338 (RPS3A-like)	113,220,613	113,221,812	MC
FCAB-89F20_Transcript1	113,531,770*	113,689,137*	SC
LOC102901637 (INTS6LB)	114,182,695	114,188,831	MC
LOC102900259 (INTS6LC)	114,222,292	114,230,935	MC
LOC102901675 (INTS6LD)	114,259,526	114,268,867	MC
LOC101083152	114,564,281	114,567,927	SC
LOC102899102 (TNP4L)	117,716,423	117,717,176	MC
LOC111558724 (FH14)	122,380,408	122,383,513	MC
LOC101091265 (ELL2)	122,482,843	122,491,048	MC
LOC102901550 (TNP4)	122,650,566	122,651,363	MC
LOC111558690 (BRP-like)	124,871,180	124,888,322	MC

Table B3.12. Continued

Gene Name	Cat Start	Cat End	Status
LOC111558732	127,862,926	127,864,162	SC
PNMA6FL	128,050,943	128,059,895	MC
LOC111558735 (MUM1L1)	128,235,772	128,237,735	MC
LOC105259690 (SLC6A8-like)	129,432,061	129,435,675	MC
FCAB-331E2-chrX_random_ctg158_Transcript	129,435,118*	129,505,645*	SC
FCAB-331E2_Transcript2	129,435,118*	129,505,645*	MC
FCAB-331E2_Transcript3	129,435,118*	129,505,645*	SC
FCAB-331E2_Transcript4	129,435,118*	129,505,645*	MC
FCAB-331E2_Transcript5	129,435,118*	129,505,645*	SC
GTPBP6AL	130,463,634	130,469,229	MC
PLCXD1L	130,473,426	130,502,832	MC

Table B3.13. Locations of novel genes on the pige X chromosome. Genes within ampliconic regions are italicized, underlined, and in bold.

Gene Name	Pig Start	Pig End	Status
LOC110257789	418,073	433,468	SC
LOC110257836	2,127,385	2,145,225	SC
LOC110257838	6,295,275	6,296,381	SC
LOC110257839	9,501,052	9,505,154	SC
CXH4orf3	12,139,795	12,141,315	MC
LOC100156694	20,366,396	20,433,395	SC
LOC102167889 (RNF168)	20,654,157	20,659,555	MC
LOC100513448 (MAGEB4L)	22,170,175	22,171,346	MC
LOC100157863 (APOPT1)	22,381,742	22,382,524	MC
LOC100514934 (RPL39L)	23,259,680	23,259,840	MC
LOC100156962 (MAGEB10L)	23,652,358	23,653,374	MC
LOC102159844 (MAGEB4L)	25,895,424	26,030,203	MC
LOC100623332 (MAGEB16L)	***31,357,792***	***31,721,181***	Ampliconic
LOC100522230 (MAGEB16L)	***31,573,305***	***31,574,593***	Ampliconic
LOC100522411 (MAGEB16L)	***31,592,308***	***31,597,919***	Ampliconic
LOC100623432 (MAGEB16L)	***31,752,126***	***31,753,755***	Ampliconic
LOC100520818 (MAGEB16L)	***31,759,602***	***31,780,300***	Ampliconic
LOC110257841 (H2A-Bbd type 1-like)	33,299,003	33,299,753	MC
LOC100624295 (H2A-Bbd type 1-like)	33,307,648	33,308,398	MC
LOC100736633 (H2A-Bbd type 1-like)	33,894,983	33,895,366	MC
LOC106506945 (SRRM1L)	35,880,111	35,904,287	MC
LOC110257671	36,042,052	36,140,080	SC
LOC110257750	37,168,973	37,172,889	SC
LOC110257673 (PRB2-like)	41,579,893	41,581,086	MC
L1TD1	46,349,921	46,352,855	MC
LOC106506967 (SMC1AL)	46,568,746	46,570,505	MC
LOC106506966	47,552,766	47,561,116	SC
LOC100511853 (GMCL1L1)	49,726,706	49,729,782	MC
LOC110257675 (GMCL1L1)	***49,845,934***	***49,893,816***	Ampliconic
LOC110257676 (GMCL1L1)	49,896,930	49,898,438	MC
LOC100621166 (CXORF49L)	57,678,534	57,682,641	MC
ENSSSCG00000031328 (DMRTC1L)	***58,717,308***	***58,719,987***	Ampliconic
LOC100525624 (NAP1L2L)	58,941,164	58,951,122	MC
LOC110257684 (DPH3 homolog)	71,948,753	71,949,100	MC
LOC100625147 (SNURF)	79,811,978	79,812,192	SC
LOC110257759	81,543,472	81,545,852	SC

Table B3.13. Continued

Gene Name	Pig Start	Pig End	Status
LOC106508237	84,409,372	84,468,158	SC
LOC110257688 (CFAP251-like)	85,039,485	85,040,165	MC
LOC100515519 (NCL-like)	86,974,106	86,976,917	MC
LOC100738042 (MIC10-like)	92,197,909	92,199,862	MC
LOC110257690	95,554,791	95,654,903	SC
LOC414400 (RPLP1L)	96,800,950	96,875,862	MC
ADH4	98,582,002	98,583,002	MC
CCDC169	106,457,036	106,457,680	MC
LOC100511059 (*OR1020L*)	107,752,365	107,762,735	MC
LOC110257712 (FAM127L)	110,813,912	110,815,239	MC
LOC110255250 (SIM10L2A)	110,872,816	110,876,723	MC
LOC100157667 (INTS6L)	111,094,016	111,110,838	MC
LOC100621767 (ADGRG4L)	111,469,540	111,488,383	MC
LOC102158723 (Porcine ERV)	112,845,408	112,848,002	MC
LOC110257694 (PRB2-like)	113,340,076	113,341,321	MC
LOC100523110 (RPS27L)	114,070,842	114,071,361	MC
LOC110257749 (SEC31L)	114,575,980	115,026,005	MC
LOC110257823	115,725,317	115,728,246	SC
LOC102157658 (MAGEA13P)	118,448,490	118,566,819	MC
LOC100511430 (TNP4)	118,621,838	118,622,598	MC
LOC110257797	121,231,431	121,500,019	SC
LOC110257715 (MAGEA10L)	***121,632,759***	***121,634,665***	Ampliconic
LOC100156285 (MAGEA10L)	121,688,350	121,691,552	MC
LOC100514940 (MAGEA10L)	121,699,620	121,702,252	MC
LOC100513838 (MAGEA10L)	121,708,847	121,709,729	MC
LOC100513259 (MAGE8L)	121,755,015	121,756,449	MC
LOC106504240 (MAGE10L)	121,763,484	121,765,954	MC
LOC106506163	121,776,859	121,800,481	SC
LOC110257698	123,869,573	123,870,559	SC
LOC100154655 (SYCP3L)	124,078,500	124,089,002	MC
LOC110257703 (PNMA6FL)	124,203,040	124,204,023	MC
LOC110257704 (COL3A1)	124,870,296	124,881,007	MC
LOC100737597 (IF-2-like)	125,072,091	125,097,974	MC
LOC102158155 (CXorf49L)	125,109,070	125,112,343	MC
LOC110257705 (OR10VL)	125,122,994	125,125,168	MC
LOC102166335 (LAGE3-like)	125,150,593	125,152,014	MC
LOC106507061 (EVA1C-like)	125,579,523	125,595,031	MC